

Local/global model order reduction strategy for the simulation of quasi-brittle fracture

P. Kerfriden¹, J.C. Passieux², S.P.A. Bordas¹

¹ Cardiff University, School of Engineering
Institute of Mechanics and Advanced Materials
Queen's Buildings, The Parade, Cardiff CF24 3AA, Wales, United Kingdom

² Université de Toulouse/INSA
Institut Clément Ader (ICA)
133 Avenue de Rangueil, F-31077 Toulouse, France

September 27, 2011

Abstract

This paper proposes a novel technique to reduce the computational burden associated with the simulation of localised failure. The proposed methodology affords the simulation of damage initiation and propagation whilst concentrating the computational effort where it is most needed, i.e. in the localisation zones. To do so, a local/global technique is devised where the global (slave) problem (far from the zones undergoing severe damage and cracking) is solved for in a reduced space computed by the classical Proper Orthogonal Decomposition, while the local (master) degrees of freedom (associated with the part of the structure where most of the damage is taking place) are fully resolved. Both domains are coupled through a local/global technique. This method circumvents the difficulties associated with model order reduction for the simulation of highly non-linear mechanical failure and offers an alternative or complementary approach to the development of multiscale fracture simulators.

Keywords: Adaptive Model Order Reduction (MOR); Local/global Approach, Nonlinear Fracture Mechanics; Proper Orthogonal Decomposition (POD); Newton/Krylov Solver;

1 Introduction

Simulating damage initiation and subsequent global structural failure is one of the most active topics in computational mechanics. Several mathematical models and numerical methods have been developed over the years to assess various limit states such as failure due to permanent deformations, cracks or decohesion/delamination, e.g. in composite materials. Yet, these models, be they damage based or relying on discrete cracks are generally computationally expensive, as they require a fine scale description of the structural and material properties. Therefore, today's engineers are not able to use these state-of-the-art models for routine design. For important recent advances in the treatment of material failure (e.g. discontinuous fracture [1], advanced damage models [2, 3], damage plasticity models [4, 5] or their combination [6], etc.) to become useful in practice, it is thus important to devise techniques which are able to significantly reduce the computational effort required without sacrificing accuracy.

Historically, reducing the computational time associated with solving nonlinear problems in solid mechanics has mainly been addressed by developing homogenisation techniques [7, 8]. In this case, the material properties associated with a material point in a coarse representation of the structure

is obtained by averaging of the fine scale material behaviour over a “representative volume element” which suppresses the need to resolve the fine scale explicitly over the (possibly large) structure. These methods are widely adopted (see e.g. for a recent description [9]) and are rather effective in decreasing the computational costs associated with solving fine scale problems.

However, the application of homogenisation-based methods to failure simulations is not straightforward, as the assumptions required to prove the separation of scales are not satisfied. This may explain why such methods are still under-developed for fracture problems. The ability to coarse-grain fracture information would however lead to a further decrease of computational time by concentrating the computational effort where it is needed most. In the wake of a propagating crack, for example, the full resolution used to resolve fine scale effects close to the crack front is no longer necessary, and a coarse description may be sufficient. This issue is being intensively investigated by several teams, namely [10] who developed a coarse graining method called Multiscale Aggregating Discontinuities, and [11, 12, 13] who were able to derive an homogenisation scheme for lumping the fine scale defects of Representative Volume Elements (RVEs) into cohesive cracks introduced at the coarse scale. The solutions proposed so far are however highly problem-dependent and usually do not inherit the solid mathematical foundations from homogenisation theory.

We propose here to follow an alternative route which relies on very recent advances in model order reduction strategies for highly nonlinear structural problems [14, 15, 16, 17, 18]. We believe that, though rarely studied in the context of failure simulation in nonlinear mechanics, this avenue provides a possibly interesting alternative (or complementary) approach to homogenisation-based solution procedures. This paper is in substance a significant improvement of the work proposed in [18], dedicated to the fast prediction of the onset of structural failure by adaptive model order reduction. We aim here at preserving the computational efficiency and accuracy obtained in [18] when the size of the damaged zone becomes significant compared to the overall structure.

The definition of model order reduction can vary depending on authors. We speak here of an approximation of the unknown field in a coarse space spanned by a basis of global vectors (as opposed to locally supported shape functions, in finite elements for example). Model order reduction (MOR) finds its roots in the systems engineering community, where the issue of obtaining a fast and reliable approximation of transfer functions has been extensively addressed over the years. The interested reader can refer to the review proposed in [19]. In the context of solid mechanics, most studies have focused on the reduction of dynamics problems. Starting from classical modal basis [20] in a Ritz-Galerkin framework, two main classes of model order reduction methods have been proposed. The first propose a truncation of the modal basis and an approximation of the high-frequency behaviour by static corrections (see for instance [21]). The second use the idea of domain partitioning. Slave degrees of freedom are eliminated by Guyan reduction (or static condensation). The famous Craig-Bampton method [22] is an evolution of this idea, using truncated modal bases to enhance static condensation. These techniques have been extended to the approximation of nonlinear dynamical behaviour (an interesting review is proposed in [23]).

The literature concerning the application of model order reduction to the approximation of the (quasi-)static behaviour of nonlinear structures is a lot sparser. Different routes have been investigated to develop such reduced models. The first family of methods consists in taking advantage of the a priori knowledge of a set of representative solutions to fasten the solution process. These representative responses, called snapshots [24], can be for instance a set of solution vectors obtained for given values of structural parameters. The Proper Orthogonal Decomposition (POD) [25, 26] can classically be used to extract a relevant global basis from these snapshot vectors, which in turn defines a coarse space where the solution to the structural problem with a different set of parameters can be cheaply searched. The second family is not exactly based on reduction by projection. Spectral approaches referred to as Proper Generalized Decomposition (PGD) are a family of solvers capable of building the solution to a multidimensional reference problem as a sum of products of functions of separate variables. The dimensions can be space or time, but also structural parameters. Once this solution is obtained *offline*, the *online* computation simply consists in the evaluation of the separated functions, which are known explicitly. These novel methods are very appealing and were recently applied to linear, stochastic

and/or time dependant problems [27, 28] (including linear predictions of nonlinear solvers [29]), and to non-linear partial differential equations [30, 28]. Their extension to irreversible constitutive laws such as the ones considered in damage mechanics does not seem to be straightforward, and, as far as we know, has never been addressed before (it is therefore a very interesting possible extension of the present piece of work).

We focus here on the first family of solvers. To the authors' knowledge, the application of model order reduction methods to failure simulations has virtually never been addressed. The main reason for this is that the solution fields in the zones of the structure where damage or failure initiate and propagate are highly parameter-dependent, and cannot be known in advance. Hence, precomputed reduced bases are unable to represent the structural behaviour beyond the onset of strain localisation, damage and fracture. In the particular case of linear elastic fracture mechanics, the authors of [31] have proposed a mesh morphing technique that allows to keep an autosimilar description of the singularity during propagation. In this particular case, the use of a snapshot POD is relevant and allow for interesting computational savings. Corrective POD algorithms have also been proposed to alleviate this in the case of plasticity or damage models [15, 18]. Both these techniques are based on global corrections of the initial reduced basis by Krylov algorithms. Though efficient for global nonlinearities, it was shown in [18] that in the case of damage assessment, the number of corrections increases when nonlinearities localise, i.e.: when failure occurs. The purpose of this paper is to circumvent this phenomenon.

The first stage of the proposed method is to only reduce the set of balance equations which exhibits a "smooth" nonlinearity. In the context of the simulation of damage, one part of the degrees of freedom, corresponding to highly damaged zones, will be considered as master degrees of freedom, which will be fully solved for, while the remaining will be approximated as a linear combination of reduced basis vectors, which may be precomputed and corrected on the fly. In this respect, the proposed method has strong links with Craig-Bampton methods, or model order reduction by substructure such as the one proposed in [32, 33, 29]. A similar approach was also proposed in [34]. The second stage of our developments consists in developing an efficient solver for the linearized system of equations obtained by a Newton solution algorithm for the reduced problem. The system is condensed on the master degrees of freedom, and solved by a Krylov algorithm. The global reduced basis is used to efficiently solve the condensed problem: an augmentation technique ensures that the iterations are performed in a space which is orthogonal to the one spanned by the reduced basis. As a result, the augmented Krylov iterations correspond to a two-level solver, the coarse level being the space spanned by the restriction of reduced basis vectors to the fully resolved set of degrees of freedom.

The paper is organized as follows. In the first section, we recall the basics of projection-based model order reduction applied to the solution of nonlinear quasi-static problems. In section 3, we formally describe the proposed local/global reduction algorithm. This technique is applied to the reduction of damage models in section 4. Finally, we show in the last section that this technique permits to drastically reduce the number of corrections of the reduced basis required to obtain a desired level of accuracy.

2 Model order reduction by projection: general concepts

Projection-based model order reduction can be introduced as a means to obtain, at low computational cost, a good approximation to a set of nonlinear equations of the form:

$$\underline{\mathbf{F}}\left(\left(\underline{\mathbf{X}}_i\right)_{i \in \llbracket 0, N \rrbracket}\right) = \underline{\mathbf{0}} \quad (1)$$

Where $\underline{\mathbf{F}}$ is a vectorial, possibly nonlinear, function of a set of vectorial state variables $\underline{\mathbf{X}}$. The expression of this function can take various forms, depending on the application.

2.1 Problem statement in nonlinear continuum mechanics

We consider a structure occupying a continuous domain Ω with boundary $\partial\Omega$. It is subjected to prescribed displacements \underline{U}_D on its boundary $\partial\Omega_u$ and to prescribed tractions \underline{F}_D on the complementary boundary $\partial\Omega_f = \partial\Omega \setminus \partial\Omega_u$, over time interval $[0, T]$. Let \underline{u} be the unknown displacement field, which belongs to the space \mathcal{U} of kinematically admissible fields:

$$\mathcal{U} = \left\{ \underline{u} \in H^1(\Omega) \mid \underline{u}|_{\partial\Omega_u} = \underline{U}_D \right\} \quad (2)$$

Let \mathcal{U}^0 be the associated vector space.

$$\mathcal{U}^0 = \left\{ \underline{u} \in H^1(\Omega) \mid \underline{u}|_{\partial\Omega_u} = \underline{0} \right\} \quad (3)$$

Under the assumptions of quasi-static evolution of the structure and small perturbations, the weak form of the equilibrium and constitutive laws read, at any time $t \in [0, T]$:

$$\begin{aligned} \forall \underline{u}^* \in \mathcal{U}^0, \text{ find } \underline{u} \in \mathcal{U} \text{ such that:} \\ \int_{\Omega} \underline{\underline{\sigma}} : \underline{\underline{\epsilon}}(\underline{u}^*) \, d\Omega = \int_{\Omega} \underline{\underline{f}}_D \cdot \underline{u}^* \, d\Omega + \int_{\partial\Omega_f} \underline{F}_D \cdot \underline{u}^* \, d\Gamma \\ \underline{\underline{\sigma}} = \underline{\underline{\sigma}} \left((\underline{\underline{\epsilon}}(\underline{u}))_{|\tau} \right)_{\tau \leq t} \end{aligned} \quad (4)$$

where $\underline{\underline{\sigma}}$ is the Cauchy stress tensor and $\underline{\underline{\epsilon}}(\underline{u})$ is the symmetric part of the displacement gradient. The constitutive relation between $\underline{\underline{\sigma}}$ and $\underline{\underline{\epsilon}}(\underline{u})$ is nonlinear and described using internal variables (plasticity, damage for instance). It is assumed local and rate-independent.

2.1.1 Space and time discretization

Let us perform a standard finite element approximation of the space of unknown displacement field \mathcal{U} (and a similar approximation of the space of test functions \mathcal{U}^0):

$$\mathcal{U}^h(\Omega) = \left\{ \underline{u}(\underline{x}) \mid \underline{u}(\underline{x}) = \sum_{i=1}^{n_n} N_i(\underline{x}) \underline{u}_i \right\} \quad (5)$$

where n_n is the number of nodes, \underline{x} is the position vector, $(N_i)_{i \in [1, n_n]}$ are the finite element shape functions associated to each node of the finite element mesh, and $(\underline{u}_i)_{i \in [1, n_n]}$ are the nodal values of the displacement field.

The nonlinear solution strategy used to solve the problem over time is a classical time discretization scheme for quasi-static and rate-independent problems. This procedure consists in finding a set of consecutive solutions at times $(t_n)_{n \in [0, n_t]}$.

The introduction of the finite element approximation and time discretization into equation (4) at any time t_{n+1} of the analysis leads to the following nonlinear vectorial equation:

$$\underline{\mathbf{F}}_{\text{Int}} \left(\underline{\Delta \mathbf{U}}, \left(\underline{\mathbf{U}}|_{t_m} \right)_{m \in [0, n]} \right) + \underline{\mathbf{F}}_{\text{Ext}} = \underline{\mathbf{0}} \quad (6)$$

where the vector of increment in the nodal unknowns $\underline{\Delta \mathbf{U}} \in \mathbb{R}^{n_u}$ (n_u is the number of scalar nodal unknowns) is defined by $\underline{\Delta \mathbf{U}} = \underline{\mathbf{U}}|_{t_{n+1}} - \underline{\mathbf{U}}|_{t_n}$, $\underline{\mathbf{F}}_{\text{Int}} \in \mathbb{R}^{n_u}$ and $\underline{\mathbf{F}}_{\text{Ext}} \in \mathbb{R}^{n_u}$ are respectively the internal forces resulting from the discretization of the internal virtual work (left-hand side of the first equation of system (4)) and the external forces resulting from the discretization of the external virtual work (right-hand side of the first equation of system (4)). In the following, the dependency of the internal forces with respect to the history of the unknown fields will not be written explicitly, unless necessary.

The set of successive solution vectors to problem (6) at times $(t_n)_{n \in [0, n_t]}$, obtained by a tangent Newton algorithm, will be considered as the reference solution. The algorithms used in the following should provide a solution that is closer to the reference solution, but for a cheaper computational cost. Therefore, the relevance of (i) the initial space and time discretizations and (ii) the chosen Newton solution algorithm will not be discussed in this paper.

2.2 Reduced solution of the balance equations by projection

2.2.1 Principle

The solution vector is searched for in a space of small dimension (several orders smaller than the number of finite element degrees of freedom). Let us call $\underline{\underline{\mathbf{C}}}$ the matrix whose columns form a basis of this space:

$$\underline{\underline{\mathbf{C}}} = (\underline{\mathbf{C}}^1 \quad \underline{\mathbf{C}}^2 \quad \dots \quad \underline{\mathbf{C}}^{n_c}) \quad (7)$$

where n_c is the dimension of the reduced space, and $(\underline{\mathbf{C}}^k)_{k \in \llbracket 1, n_c \rrbracket} \in (\mathbb{R}^{n_u})^{n_c}$ are the basis vectors. Applied to the reduction of problem (6), the increment in the solution field is approximated by:

$$\underline{\Delta \mathbf{U}} = \underline{\underline{\mathbf{C}}} \underline{\alpha} \quad (8)$$

where we introduced the reduced state variables $\underline{\alpha} \in \mathbb{R}^{n_c}$. Problem (6) is now overconstrained. One usually tackle this problem by prescribing a Galerkin orthogonality condition: the residual of equation (6) $\underline{\mathbf{R}} = \underline{\mathbf{F}}_{\text{Int}}(\underline{\mathbf{U}}|_{t_n} + \underline{\underline{\mathbf{C}}} \underline{\alpha}) + \underline{\mathbf{F}}_{\text{Ext}}$ is constrained to be orthogonal to any test vector $\underline{\delta \mathbf{U}}^* = \underline{\underline{\mathbf{C}}} \underline{\delta \alpha}^*$ belonging to the space spanned by the reduced basis vectors $(\underline{\mathbf{C}}^k)_{k \in \llbracket 1, n_c \rrbracket}$:

$$\underline{\underline{\mathbf{C}}}^T \underline{\mathbf{R}} = \underline{\mathbf{0}} \quad (9)$$

The reduced form of problem (6) is:

$$\underline{\underline{\mathbf{C}}}^T \left(\underline{\mathbf{F}}_{\text{Ext}} + \underline{\mathbf{F}}_{\text{Int}}(\underline{\mathbf{U}}|_{t_n} + \underline{\underline{\mathbf{C}}} \underline{\alpha}) \right) = \underline{\mathbf{0}} \quad (10)$$

2.2.2 One particular choice of projection space: the proper orthogonal decomposition (POD)

Various choices of projection spaces are proposed in the literature. One of the most successful, as far as nonlinear simulations are concerned, is the snapshot-POD [24]. This particular technique requires the knowledge of a representative family of solutions to the global problem. This set of vectors $(\underline{\mathbf{S}}^k)_{k \in \llbracket 1, n_s \rrbracket}$ is called *snapshot*. The aim is to find an orthonormal basis $(\underline{\mathbf{C}}^k)_{k \in \llbracket 1, n_c \rrbracket}$, of dimension n_c smaller than n_s such that the distance between spaces $\text{Im}(\underline{\underline{\mathbf{C}}})$ and $\text{Im}(\underline{\underline{\mathbf{S}}})$ is minimum. This distance can be quantified by the following metric:

$$J(\underline{\mathbf{C}}^1, \dots, \underline{\mathbf{C}}^{n_c}) = \sum_{j=1}^{n_s} \left\| \underline{\mathbf{S}}^j - \sum_{i=1}^{n_c} \left(\underline{\mathbf{C}}^{iT} \underline{\mathbf{S}}^j \right) \underline{\mathbf{C}}^i \right\|^2 \quad (11)$$

which must be minimized under the constraint of orthonormality of $(\underline{\mathbf{C}}^k)_{k \in \llbracket 1, n_c \rrbracket}$. Problem (11) can be reformulated as an unconstrained formulation:

$$L(\underline{\mathbf{C}}^1, \dots, \underline{\mathbf{C}}^{n_c}, \lambda_{11}, \dots, \lambda_{n_c n_c}) = J(\underline{\mathbf{C}}^1, \dots, \underline{\mathbf{C}}^{n_c}) + \sum_{i=1}^{n_c} \sum_{j=i}^{n_c} \lambda_{ij} \left(\underline{\mathbf{C}}^{iT} \underline{\mathbf{C}}^j - \delta_{ij} \right) \quad (12)$$

where $(\lambda_{ij})_{i,j \in \llbracket 1, n_c \rrbracket}$ are Lagrange multipliers and δ_{ij} the Kronecker symbol. The $n_c(n_c+1)/2$ optimality conditions for maximizing the Lagrangian with respect to the lagrange multiplier naturally ensure the orthogonality of the basis:

$$\frac{\partial L}{\partial \lambda_{ij}} = 0 \quad \rightarrow \quad \underline{\mathbf{C}}^{iT} \underline{\mathbf{C}}^j = \delta_{ij} \quad (13)$$

Whereas the n_c optimality conditions for minimizing the Lagrangian with respect to the basis vectors reads:

$$\frac{\partial L}{\partial \underline{\mathbf{C}}^i} = \underline{\mathbf{0}} \quad \rightarrow \quad \sum_{j=1}^{n_s} \underline{\mathbf{S}}^j \left(\underline{\mathbf{C}}^{iT} \underline{\mathbf{S}}^j \right) = \lambda_{ii} \underline{\mathbf{C}}^i \quad \text{and} \quad \lambda_{ij} = 0 \quad \text{for} \quad i \neq j \quad (14)$$

Defining $\lambda_i = \lambda_{ii}$ and $\underline{\underline{\mathbf{S}}} = (\underline{\mathbf{S}}^1 \quad \underline{\mathbf{S}}^2 \quad \dots \quad \underline{\mathbf{S}}^{n_s}) \in \mathbb{R}^{n_u} \times \mathbb{R}^{n_s}$, equations (14) can be written as the following $n_u \times n_u$ eigenvalue problem:

$$\underline{\underline{\mathbf{S}}} \underline{\underline{\mathbf{S}}}^T \underline{\mathbf{C}}^i = \lambda_i \underline{\mathbf{C}}^i \quad (15)$$

This problem is equivalent to computing the Singular Value Decomposition (SVD) of $\underline{\underline{\mathbf{S}}}$, where the singular values s_i are such that $s_i^2 = \lambda_i \geq 0$. If $n_s \ll n_u$ (which is the case in general when applying the snapshot-POD), replacing problem (15) by the eigenvalue problem on the $n_s \times n_s$ operator $\underline{\underline{\mathbf{S}}}^T \underline{\underline{\mathbf{S}}}$ is computationally cheaper, and reads:

$$\underline{\mathbf{C}}^i = \lambda_i^{-\frac{1}{2}} \underline{\underline{\mathbf{S}}} \underline{\mathbf{V}}^i \quad \text{with} \quad \underline{\underline{\mathbf{S}}}^T \underline{\underline{\mathbf{S}}} \underline{\mathbf{V}}^i = \lambda_i \underline{\mathbf{V}}^i \quad (16)$$

The error associated with the truncation of the SVD at rank n_c is quantified by the distance $J(\underline{\mathbf{C}}^1, \dots, \underline{\mathbf{C}}^{n_c})$ which can be shown to be equal to the sum of the truncated eigenvalues:

$$J(\underline{\mathbf{C}}^1, \dots, \underline{\mathbf{C}}^{n_c}) = \sum_{i=n_c+1}^{n_s} \lambda_i \quad (17)$$

This result provides a precious tool for choosing the most relevant basis. Indeed, only the basis vectors $\underline{\mathbf{C}}_i$ associated with the largest λ_i (in practice, those such that $\lambda_i/\lambda_{max} > \varepsilon$, ε being a small parameter depending on the required accuracy) are selected as reduced basis vectors. This ensures that for an given number n_c of basis vectors, the approximation of the snapshot space is maximised in the sense of (11).

2.3 Nonlinear solution algorithm

At any time step, problem (10) can be solved by classical Newton algorithms (successive linearizations). Iteration $i + 1$ of the tangent Newton method consists in solving the following linear predictor:

$$\underline{\underline{\mathbf{K}}}_{\text{T,R}}^i \underline{\delta \alpha}^{i+1} = -\underline{\mathbf{R}}_{\text{R}}^i \quad (18)$$

where the increment in the reduced state variables is $\underline{\delta \alpha}^{i+1} = \underline{\alpha}^{i+1} - \underline{\alpha}^i$, and the residual and tangent operator obtained from the knowledge of the solution at iteration i read (see [18] for more details):

$$\begin{cases} \underline{\mathbf{R}}_{\text{R}}^i = \underline{\mathbf{C}}^T \left(\underline{\mathbf{F}}_{\text{Int}}(\underline{\mathbf{U}}|_{t_n} + \underline{\mathbf{C}} \underline{\alpha}^i) + \underline{\mathbf{F}}_{\text{Ext}} \right) \\ \underline{\underline{\mathbf{K}}}_{\text{T,R}}^i = \underline{\mathbf{C}}^T \left. \frac{\partial \underline{\mathbf{F}}_{\text{Int}}(\underline{\mathbf{U}}|_{t_n} + \underline{\mathbf{C}} \underline{\alpha})}{\partial \underline{\alpha}} \right|_{\underline{\alpha}=\underline{\alpha}^i} \end{cases} \quad (19)$$

3 Local/global model order reduction strategy

Projection-based model order reduction techniques introduce a global approximation of the displacement. As such, even adaptive versions [35, 18] of these methods are not well suited to the analysis of localised nonlinearities. We propose in the following to use reduction techniques by projection for the “weakly nonlinear” equations of reference problem (6), namely equations for which reduced order modelling (ROM) is relevant, while the remaining equations will be solved directly.

3.1 Displacement approximation

The principle of the proposed local/global strategy is to split the unknown solution vector into two parts, only one of them being approximated as a linear combination of reduced basis vectors. We shall use superscript (r) for the approximated “slave” part of the solution vector, while superscript (f) (which stands for “fully resolved”) will be used to denote its complementary “master” part.

More precisely, let $\widetilde{\underline{\Delta \mathbf{U}}}$ be the unknown increment vector at time t_{n+1} , the numbering being reorganised as follows:

$$\widetilde{\underline{\Delta \mathbf{U}}} = \begin{pmatrix} \widetilde{\underline{\Delta \mathbf{U}}}^{(r)} \\ \widetilde{\underline{\Delta \mathbf{U}}}^{(f)} \end{pmatrix} = \begin{pmatrix} \underline{\mathbf{E}}^{(r)} \\ \underline{\mathbf{E}}^{(f)} \end{pmatrix} \underline{\Delta \mathbf{U}} \quad (20)$$

In the above notations, $\widetilde{\underline{\Delta \mathbf{U}}}^{(r)} \in \mathbb{R}^{n_r}$ and $\widetilde{\underline{\Delta \mathbf{U}}}^{(f)} \in \mathbb{R}^{n_f}$, with $n_r + n_f = n_u$ and n_f much smaller than n_r . $\underline{\mathbf{E}}^{(r)} \in \{0, 1\}^{n_r} \times \{0, 1\}^{n_u}$ and $\underline{\mathbf{E}}^{(f)} \in \{0, 1\}^{n_f} \times \{0, 1\}^{n_u}$ are two boolean extractors (rectangular matrices with one 1 per line, the other coefficients being null).

$\widetilde{\underline{\Delta \mathbf{U}}}^{(r)}$ is approximated as a linear combination of global vectors $(\underline{\mathbf{C}}^k)_{k \in \llbracket 1, n_c \rrbracket} \in (\mathbb{R}^{n_u})^{n_c}$:

$$\widetilde{\underline{\Delta \mathbf{U}}}^{(r)} = \left(\underline{\mathbf{E}}^{(r)} \underline{\mathbf{C}} \right) \underline{\alpha} \quad (21)$$

where $\underline{\alpha} \in \mathbb{R}^{n_c}$ is the vector of reduced degrees of freedom. In the initial numbering, the unknown vector can be expressed in the form:

$$\begin{aligned} \underline{\Delta \mathbf{U}} &= \underline{\Delta \mathbf{U}}^{(r)} + \underline{\Delta \mathbf{U}}^{(f)} \\ \text{where } \begin{cases} \underline{\Delta \mathbf{U}}^{(r)} &= \underline{\mathbf{E}}^{(r)T} \widetilde{\underline{\Delta \mathbf{U}}}^{(r)} \\ \underline{\Delta \mathbf{U}}^{(f)} &= \underline{\mathbf{E}}^{(f)T} \widetilde{\underline{\Delta \mathbf{U}}}^{(f)} \end{cases} \end{aligned} \quad (22)$$

Introducing projector $\underline{\mathbf{P}}^{(r)} = \underline{\mathbf{E}}^{(r)T} \underline{\mathbf{E}}^{(r)}$, the approximation of the displacement increment finally reads:

$$\underline{\Delta \mathbf{U}} = \underline{\mathbf{P}}^{(r)} \underline{\mathbf{C}} \underline{\alpha} + \underline{\mathbf{E}}^{(f)T} \widetilde{\underline{\Delta \mathbf{U}}}^{(f)} \quad (23)$$

For the sake of legibility, we define the new state vector $\underline{\mathbf{X}}$ at time t_{n+1} :

$$\underline{\mathbf{X}} = \begin{pmatrix} \underline{\alpha} \\ \widetilde{\underline{\Delta \mathbf{U}}}^{(f)} \end{pmatrix} \quad (24)$$

Using the above definition, the approximation of the displacement increment is now:

$$\underline{\Delta \mathbf{U}} = \underline{\mathbf{A}} \underline{\mathbf{X}} \quad \text{where} \quad \underline{\mathbf{A}} = \begin{pmatrix} \underline{\mathbf{P}}^{(r)} \underline{\mathbf{C}} & \underline{\mathbf{E}}^{(f)T} \end{pmatrix} \quad (25)$$

3.2 Locally reduced set of balance equations

At that point, we need to define how the overconstrained balance equations (6) will be solved for. Let us restrict our numerical investigations to the Galerkin procedure. The balance equation are therefore required to be orthogonal to any test vector:

$$\underline{\delta \mathbf{U}}^* = \underline{\mathbf{A}} \begin{pmatrix} \underline{\alpha}^* \\ \widetilde{\underline{\Delta \mathbf{U}}}^{(f)*} \end{pmatrix} \quad (26)$$

where $\underline{\delta \alpha}^* \in \mathbb{R}^{n_c}$ and $\widetilde{\underline{\delta \mathbf{U}}}^{(f)*} \in \mathbb{R}^{n_f}$ are arbitrary test state variables. The orthogonality requirement applied to the initial nonlinear set of equations (6) yields the following nonlinear system:

$$\underline{\mathbf{R}}_{\mathbf{R}}(\underline{\mathbf{X}}) = \underline{\mathbf{A}}^T \left(\underline{\mathbf{F}}_{\text{Int}} \left(\underline{\Delta \mathbf{U}}(\underline{\mathbf{X}}) + \underline{\mathbf{U}}|_{t_n} \right) + \underline{\mathbf{F}}_{\text{Ext}} \right) = \underline{\mathbf{0}} \quad (27)$$

3.3 Newton solution scheme

At each time step of the time discretization scheme, the reduced problem (27) is solved by a Newton algorithm. The $(i+1)^{th}$ Newton iteration consists in finding $\underline{\delta \mathbf{X}}^{i+1}$ satisfying the following set of $n_c + n_f$ linearized equation:

$$\left. \frac{\partial \underline{\mathbf{R}}_R(\underline{\mathbf{X}})}{\partial \underline{\mathbf{X}}} \right|_{\underline{\mathbf{X}}=\underline{\mathbf{X}}^i} \underline{\delta \mathbf{X}}^{i+1} = -\underline{\mathbf{R}}_R^i \quad (28)$$

where $\underline{\delta \mathbf{X}}^{i+1} = \underline{\mathbf{X}}^{i+1} - \underline{\mathbf{X}}^i$ and $\underline{\mathbf{R}}_R^i$ is the residual of problem (27) computed at iteration i :

$$\underline{\mathbf{R}}_R^i = \underline{\mathbf{R}}_R(\underline{\mathbf{X}}^i) \quad (29)$$

The expression of the tangent operator $\underline{\mathbf{K}}_{T,R}^i = \left. \frac{\partial \underline{\mathbf{R}}_R(\underline{\mathbf{X}})}{\partial \underline{\mathbf{X}}} \right|_{\underline{\mathbf{X}}^i}$ is obtained by differentiation of equation (27) with respect to the set of reduced state variables, in any arbitrary direction $\underline{\delta \mathbf{X}}$ such that:

$$\underline{\delta \mathbf{X}} = \begin{pmatrix} \underline{\delta \alpha} \\ \widetilde{\underline{\delta \mathbf{U}}}^{(f)} \end{pmatrix} \quad (30)$$

This differentiation reads:

$$\underline{\mathbf{K}}_{T,R}^i \underline{\delta \mathbf{X}} = \left. \frac{\partial \underline{\mathbf{R}}_R(\underline{\mathbf{X}})}{\partial \underline{\mathbf{X}}} \right|_{\underline{\mathbf{X}}^i} \underline{\delta \mathbf{X}} \quad (31)$$

Using the tangent operator of the full system of equations (6), and assuming that $\frac{\partial \underline{\mathbf{F}}_{\text{Ext}}}{\partial \underline{\Delta \mathbf{U}}} = \underline{\mathbf{0}}$, one can write:

$$\underline{\mathbf{K}}_{T,R}^i \underline{\delta \mathbf{X}} = \underline{\mathbf{A}}^T \left. \frac{\partial \underline{\mathbf{F}}_{\text{Int}}(\underline{\Delta \mathbf{U}})}{\partial \underline{\Delta \mathbf{U}}} \right|_{\underline{\Delta \mathbf{U}}=\underline{\mathbf{A}} \underline{\mathbf{X}}^i} \left. \frac{\partial \underline{\Delta \mathbf{U}}}{\partial \underline{\mathbf{X}}} \right|_{\underline{\Delta \mathbf{U}}=\underline{\mathbf{A}} \underline{\mathbf{X}}^i} \underline{\delta \mathbf{X}} \quad (32)$$

Using the notation $\underline{\mathbf{K}}_T^i = \left. \frac{\partial \underline{\mathbf{F}}_{\text{Int}}}{\partial \underline{\Delta \mathbf{U}}} \right|_{\underline{\Delta \mathbf{U}}(\underline{\mathbf{X}}^i)}$, the previous equation yields the expression of the reduced tangent operator as a function of the tangent of the full system of equations:

$$\underline{\mathbf{K}}_{T,R}^i \underline{\delta \mathbf{X}} = \underline{\mathbf{A}}^T \underline{\mathbf{K}}_T^i \underline{\mathbf{A}} \underline{\delta \mathbf{X}} \quad (33)$$

Introducing the above expression of the tangent into equation (28), and expanding the expression of operator $\underline{\mathbf{A}}$, one finally obtains the following linearized system:

$$\begin{pmatrix} \underline{\mathbf{C}}^T \underline{\mathbf{P}}^{(r)} \underline{\mathbf{K}}_T^i \underline{\mathbf{P}}^{(r)} \underline{\mathbf{C}} & \underline{\mathbf{C}}^T \underline{\mathbf{P}}^{(r)} \underline{\mathbf{K}}_T^i \underline{\mathbf{E}}^{(f)T} \\ \underline{\mathbf{E}}^{(f)} \underline{\mathbf{K}}_T^i \underline{\mathbf{P}}^{(r)} \underline{\mathbf{C}} & \underline{\mathbf{E}}^{(f)} \underline{\mathbf{K}}_T^i \underline{\mathbf{E}}^{(f)T} \end{pmatrix} \underline{\delta \mathbf{X}}^{i+1} \\ = - \begin{pmatrix} \underline{\mathbf{C}}^T \underline{\mathbf{P}}^{(r)} \\ \underline{\mathbf{E}}^{(f)} \end{pmatrix} \left(\underline{\mathbf{F}}_{\text{Int}}(\underline{\Delta \mathbf{U}}(\underline{\mathbf{X}}^i) + \underline{\mathbf{U}}|_{t_n}) + \underline{\mathbf{F}}_{\text{Ext}} \right) \quad (34)$$

Notice that problem (34) is considerably reduced compared to a direct linearisation of (6). Indeed, $\underline{\mathbf{K}}_{T,R}^i \in \mathbb{R}^{n_c} \times \mathbb{R}^{n_f}$ where $n_c \ll n_u$ and, applied to the analysis of localised nonlinearities, we can reasonably expect that n_f is at least one order of magnitude smaller than n_u , which means that only a few of the balance equations are strongly nonlinear. This assumption will be validated in section 4.

This framework is very general. It does not rely on domain decomposition but on a splitting of the balance equations. Yet, this technique will be particularized in section 4, and we will show that using the ideas inspired from domain decomposition methods [36, 37] and local/global methods [38, 39] provides efficient splitting in the case of damage mechanics.

3.4 Two-level solution of the linearized systems

The successive solution to (28) could be obtained by a direct solver. This would not affect the results given in the following. However, we believe that, for this particular application, an iterative solver has several advantages:

- Iterative solvers, such as a conjugate gradient (see [40] for instance), do not require any factorisation. Factorisations usually deteriorate the sparsity of algebraic systems, which results in an increase in the memory requirements. Newton-Krylov methods do not, in theory [41], even require the assembly of the stiffness operator.
- Direct solvers are less versatile than iterative ones, and are usually a limiting point to extend numerical solution schemes to parallel computing.
- Solving exactly the successive linearized systems (28) is not required. This idea has led to intensive work on Inexact Newton Methods [42], but this issue will not be addressed in this paper. The interested reader can refer to [37, 18] for studies on similar topics in the scope of domain decomposition and model order reduction.
- More importantly, a basic direct solver such as Cholesky factorizations will not make use of any mechanically relevant information obtained during the previous iterations of the nonlinear solver. In the chosen iterative algorithm, the part of the displacement field that is “fully resolved” will be solved for efficiently using the reduced basis as a preconditioner, following the ideas proposed for instance in [43, 41, 37].

We use here a Krylov algorithm to efficiently find a solution to linearized system of equation (28). Three stages of preconditioning are performed: a condensation, a projection and a crude preconditioning of the resulting system.

3.4.1 Condensation

Let us condense the linearized problem (34) on the master (i.e.: “fully resolved”) degrees of freedom:

$$\underline{\underline{\mathbf{S}}}_P^{(f)} \widetilde{\underline{\underline{\delta \mathbf{U}}}}^{(f)} = \underline{\underline{\mathbf{R}}}_C^{(f)} \quad (35)$$

where the superscript corresponding to the current and previous Newton iterations have been dropped for the sake of legibility. $\underline{\underline{\mathbf{S}}}_P^{(f)}$ is the primal Schur complement and $\underline{\underline{\mathbf{R}}}_C^{(f)}$ the condensed residual. They are defined by:

$$\begin{cases} \underline{\underline{\mathbf{S}}}_P^{(f)} = \underline{\underline{\mathbf{K}}}_{T,R}^{(ff)} - \underline{\underline{\mathbf{K}}}_{T,R}^{(fr)} \left(\underline{\underline{\mathbf{K}}}_{T,R}^{(rr)} \right)^{-1} \underline{\underline{\mathbf{K}}}_{T,R}^{(rf)} \\ \underline{\underline{\mathbf{R}}}_C^{(f)} = \underline{\underline{\mathbf{K}}}_{T,R}^{(fr)} \left(\underline{\underline{\mathbf{K}}}_{T,R}^{(rr)} \right)^{-1} \underline{\underline{\mathbf{R}}}_R^{(r)} - \underline{\underline{\mathbf{R}}}_R^{(f)} \end{cases} \quad (36)$$

Where we have used the usual block notations:

$$\underline{\underline{\mathbf{K}}}_{T,R} = \begin{pmatrix} \underline{\underline{\mathbf{K}}}_{T,R}^{(rr)} & \underline{\underline{\mathbf{K}}}_{T,R}^{(rf)} \\ \underline{\underline{\mathbf{K}}}_{T,R}^{(fr)} & \underline{\underline{\mathbf{K}}}_{T,R}^{(ff)} \end{pmatrix} \quad (37)$$

$$\underline{\underline{\mathbf{R}}}_R = \begin{pmatrix} \underline{\underline{\mathbf{R}}}_R^{(r)} \\ \underline{\underline{\mathbf{R}}}_R^{(f)} \end{pmatrix} \quad (38)$$

The state variables corresponding to the reduced part of the displacement vector increment have been eliminated from the system of equations and can be retrieved as follows:

$$\underline{\underline{\delta \alpha}} = - \left(\underline{\underline{\mathbf{K}}}_{T,R}^{(rr)} \right)^{-1} \left(\underline{\underline{\mathbf{R}}}_R^{(r)} + \underline{\underline{\mathbf{K}}}_{T,R}^{(rf)} \widetilde{\underline{\underline{\delta \mathbf{U}}}}^{(f)} \right) \quad (39)$$

$\left(\underline{\underline{\mathbf{K}}}_{\text{T,R}}^{(rr)}\right)^{-1}$ is an operator of very small size, which can be explicitly computed prior to applying the iterative Krylov solver. However, in the case where a very large number of reduced basis vectors is used to approximate the unknown field in the smooth region, computing the Schur-complement/vector products corresponding to the construction of the successive Krylov basis vectors will become computationally more efficient.

3.4.2 Iterative solution to the condensed problem

The classical projected (or augmented) conjugate gradient [44, 36] is applied to the approximate solution of (35) (the linearized operator is assumed symmetric, positive and definite). The chosen augmentation space is the space spanned by the reduced basis vectors $\text{Im}(\underline{\underline{\mathbf{C}}}^{(f)})$ with the restriction $\underline{\underline{\mathbf{C}}}^{(f)} = \underline{\underline{\mathbf{E}}}^{(f)} \underline{\underline{\mathbf{C}}}$.

The starting point of augmented Krylov solvers is to decompose the unknown solution increment into two supplementary spaces:

$$\begin{aligned} \widetilde{\underline{\underline{\mathbf{U}}}}^{(f)} &= \underline{\underline{\delta\mathbf{U}}}_{\text{C}}^{(f)} + \underline{\underline{\delta\mathbf{U}}}_{\text{K}}^{(f)} \\ \text{where } \begin{cases} \underline{\underline{\delta\mathbf{U}}}_{\text{C}}^{(f)} \in \text{Im}(\underline{\underline{\mathbf{C}}}^{(f)}) \\ \underline{\underline{\delta\mathbf{U}}}_{\text{K}}^{(f)} \in \text{Im}(\underline{\underline{\mathbf{C}}}^{(f)})^{\perp_{\underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)}}} = \text{Ker}(\underline{\underline{\mathbf{C}}}^{(f)T} \underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)}) \end{cases} \end{aligned} \quad (40)$$

$\perp_{\underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)}}$ designing the $\underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)}$ -orthogonality, which is ensured by introducing the classical projector:

$$\underline{\underline{\mathbf{P}}} = \underline{\underline{\mathbf{I}}}_{\text{d}} - \underline{\underline{\mathbf{C}}}^{(f)} (\underline{\underline{\mathbf{C}}}^{(f)T} \underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)} \underline{\underline{\mathbf{C}}}^{(f)})^{-1} \underline{\underline{\mathbf{C}}}^{(f)T} \underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)} \quad (41)$$

and writing that $\underline{\underline{\delta\mathbf{U}}}_{\text{K}}^{(f)} = \underline{\underline{\mathbf{P}}} \underline{\underline{\delta\mathbf{U}}}_{\text{K}}^{(f)}$.

This separation of the search space into two subspaces $\text{Im}(\underline{\underline{\mathbf{C}}}^{(f)})$ and $\text{Im}(\underline{\underline{\mathbf{P}}})$ in direct sum leads to the following uncoupled equations:

$$\begin{aligned} \underline{\underline{\delta\mathbf{U}}}_{\text{C}}^{(f)} &= \underline{\underline{\mathbf{C}}}^{(f)} (\underline{\underline{\mathbf{C}}}^{(f)T} \underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)} \underline{\underline{\mathbf{C}}}^{(f)})^{-1} \underline{\underline{\mathbf{C}}}^{(f)T} \underline{\underline{\mathbf{R}}}_{\text{C}}^{(f)} \\ \left(\underline{\underline{\mathbf{P}}}^T \underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)} \underline{\underline{\mathbf{P}}} \right) \underline{\underline{\delta\mathbf{U}}}_{\text{K}}^{(f)} &= \underline{\underline{\mathbf{P}}}^T \underline{\underline{\mathbf{R}}}_{\text{C}}^{(f)} \end{aligned} \quad (42)$$

The first line of equation (42) is a coarse initialization of the projected conjugate gradient. The second line is the linear prediction of the full problem projected on $\text{Im}(\underline{\underline{\mathbf{C}}}^{(f)})^{\perp_{\underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)}}}$. This system is symmetric and can be solved by a preconditioned conjugate gradient. Hence one solves iteratively:

$$\widetilde{\underline{\underline{\mathbf{M}}}}^{-1} \left(\underline{\underline{\mathbf{P}}}^T \underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)} \underline{\underline{\mathbf{P}}} \right) \underline{\underline{\delta\mathbf{U}}}_{\text{K}}^{(f)} = \widetilde{\underline{\underline{\mathbf{M}}}}^{-1} \underline{\underline{\mathbf{P}}}^T \underline{\underline{\mathbf{R}}}_{\text{C}}^{(f)} \quad (43)$$

where $\widetilde{\underline{\underline{\mathbf{M}}}}^{-1}$ is the left preconditioner (symmetric, definite and positive). In our test cases, $\widetilde{\underline{\underline{\mathbf{M}}}}$ is a diagonal matrix whose entries are the elements of the diagonal of $\underline{\underline{\mathbf{S}}}_{\text{P}}^{(f)}$. Algorithm 1 gives the classical implementation of an augmented preconditioned conjugate gradient.

3.4.3 Interpretation and comments on the expected additional costs

The three preconditioning techniques that we have detailed in the previous sections define a cheap structural preconditioner. A schematic representation of this algorithm is given in figure (1).

From an algebraic point of view, it can easily be proved that the condensation has the same effect on the conjugate gradient iterations as a left block preconditioner for system (34):

$$\widetilde{\underline{\underline{\mathbf{M}}}}_{\text{Condens}}^{-1} = \begin{pmatrix} \left(\underline{\underline{\mathbf{K}}}_{\text{T,R}}^{(rr)} \right)^{-1} & \underline{\underline{\mathbf{0}}} \\ \underline{\underline{\mathbf{0}}} & \underline{\underline{\mathbf{0}}} \end{pmatrix} \quad (44)$$

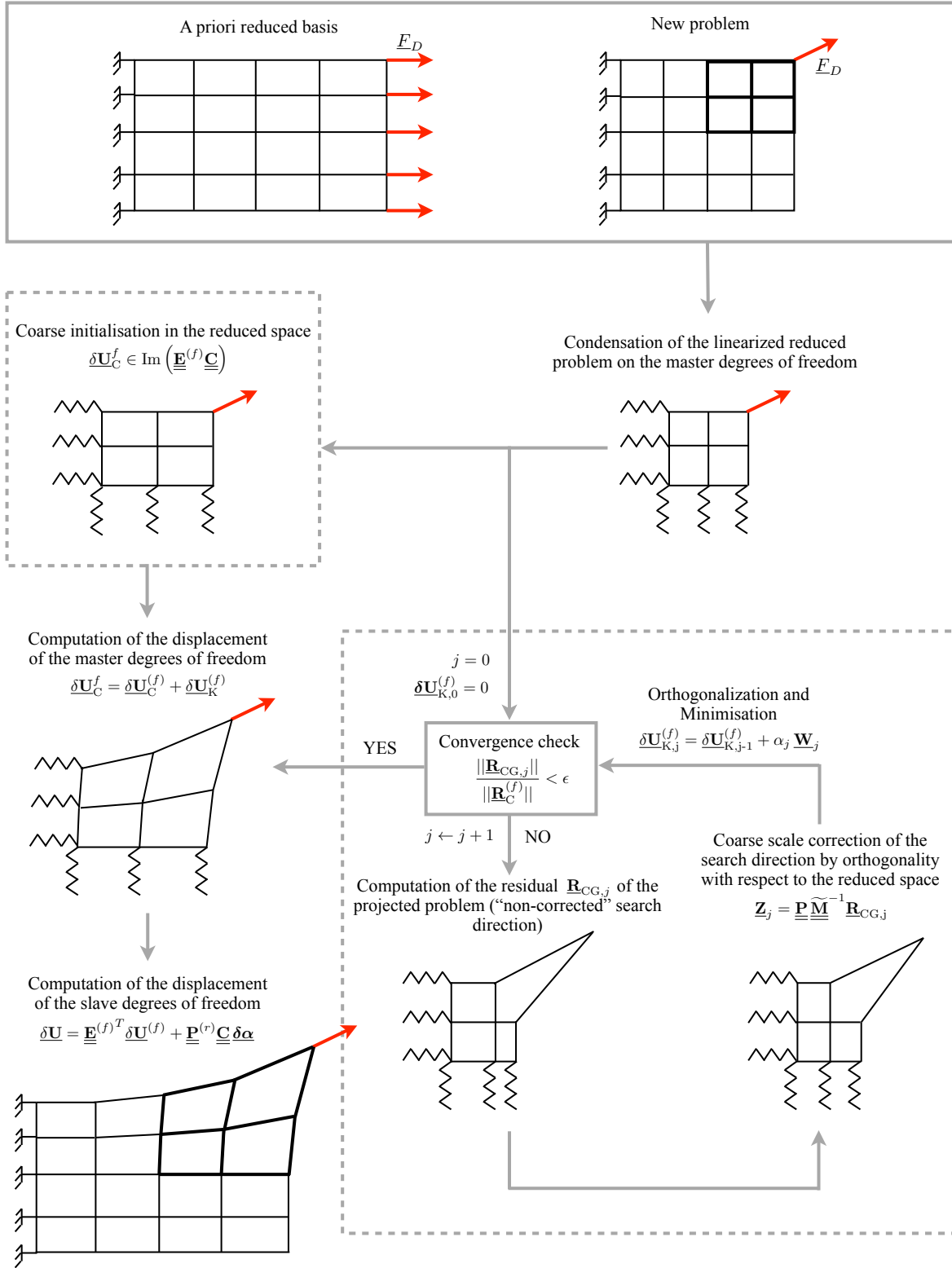


Figure 1: Graphical representation of the two-level solution algorithm.

Algorithm 1: Augmented preconditioned conjugate gradient for the solution to the condensed linearized reduced problem

Compute $\begin{cases} \underline{\mathbf{S}}_{\text{P}}^{(f)} = \underline{\mathbf{K}}_{\text{T,R}}^{(ff)} - \underline{\mathbf{K}}_{\text{T,R}}^{(fr)} \left(\underline{\mathbf{K}}_{\text{T,R}}^{(rr)} \right)^{-1} \underline{\mathbf{K}}_{\text{T,R}}^{(rf)} \\ \underline{\mathbf{R}}_{\text{C}}^{(f)} = \underline{\mathbf{R}}_{\text{R}}^{(f)} - \underline{\mathbf{K}}_{\text{T,R}}^{(fr)} \left(\underline{\mathbf{K}}_{\text{T,R}}^{(rr)} \right)^{-1} \underline{\mathbf{R}}_{\text{R}}^{(r)} \\ \underline{\mathbf{C}}^{(f)} = \underline{\mathbf{E}}^{(f)} \underline{\mathbf{C}} \\ (\underline{\mathbf{C}}^{(f)})^T \underline{\mathbf{S}}_{\text{P}}^{(f)} \underline{\mathbf{C}}^{(f)} \end{cases}^{-1}$
 $\delta \underline{\mathbf{U}}_{\text{C}}^{(f)} = \underline{\mathbf{C}}^{(f)} (\underline{\mathbf{C}}^{(f)})^T \underline{\mathbf{S}}_{\text{P}}^{(f)} \underline{\mathbf{C}}^{(f)} \underline{\mathbf{C}}^{(f)} \underline{\mathbf{R}}_{\text{C}}^{(f)} ;$
 $\underline{\mathbf{R}}_{\text{CG},0} = \underline{\mathbf{R}}_{\text{C}}^{(f)} - \underline{\mathbf{S}}_{\text{P}}^{(f)} \delta \underline{\mathbf{U}}_{\text{C}}^{(f)} = \underline{\mathbf{P}}^T \underline{\mathbf{R}}_{\text{C}}^{(f)} ;$
 $\underline{\mathbf{Z}}_0 = \underline{\mathbf{P}} \widetilde{\underline{\mathbf{M}}}^{-1} \underline{\mathbf{R}}_{\text{CG},0} \quad , \quad \underline{\mathbf{W}}_0 = \underline{\mathbf{Z}}_0 \quad \text{and} \quad \delta \underline{\mathbf{U}}_{\text{K},0}^{(f)} = 0 ;$
for $j = 1, \dots, n$ **do**
 $\alpha_{j-1} = (\underline{\mathbf{R}}_{\text{CG},j-1}, \underline{\mathbf{W}}_{j-1}) / (\underline{\mathbf{S}}_{\text{P}}^{(f)} \underline{\mathbf{W}}_{j-1}, \underline{\mathbf{W}}_{j-1})$
 $\delta \underline{\mathbf{U}}_{\text{K},j}^{(f)} = \delta \underline{\mathbf{U}}_{\text{K},j-1}^{(f)} + \alpha_{j-1} \underline{\mathbf{W}}_{j-1}$
 $\underline{\mathbf{R}}_{\text{CG},j} = \underline{\mathbf{R}}_{\text{CG},j-1} - \alpha_{j-1} \underline{\mathbf{S}}_{\text{P}}^{(f)} \underline{\mathbf{W}}_{j-1}$
 $\underline{\mathbf{Z}}_j = \underline{\mathbf{P}} \widetilde{\underline{\mathbf{M}}}^{-1} \underline{\mathbf{R}}_{\text{CG},j}$
 $\beta_j = (\underline{\mathbf{S}}_{\text{P}}^{(f)} \underline{\mathbf{W}}_{j-1}, \underline{\mathbf{Z}}_j) / (\underline{\mathbf{W}}_{j-1}, \underline{\mathbf{S}}_{\text{P}}^{(f)} \underline{\mathbf{W}}_{j-1})$
 $\underline{\mathbf{W}}_j = \underline{\mathbf{Z}}_j - \beta_j \underline{\mathbf{W}}_{j-1}$
end

From a mechanical point of view, condensing the problem on the master degrees of freedom ensures that any solution provided by the iterative solver exactly solves the reduced part of the balance equations.

The augmentation of the conjugate gradient can also be seen as a left preconditioner for the condensed system of equations (second line of system (42)), which is associated with an initialization of the iterative algorithm. The initialization $\delta \underline{\mathbf{U}}_{\text{C}}^{(f)}$ (first line of system (42)) is the best solution of the condensed problem in the restriction of the pre-computed reduced basis to the master degrees of freedom (in the sense of a $(\underline{\mathbf{S}}_{\text{P}}^{(f)})^{-1}$ -norm). Taking into account the previous comments on the interpretation of the condensation step, this initialisation is exactly the solution to the linearized system of equation (18) on the master degrees of freedom. In other words, the initialization of the iterative algorithm provides the solution of the reduced linearized problem which would have been obtained without local/global enhancement. No significant additional costs compared to the basic projection-based reduced order modelling are involved so far. In the case where the solution which is looked for is correctly approximated in the reduced space (e.g.: “smooth” nonlinearity, or unloading of the structure), no conjugate gradient iteration is required.

If conjugate gradient iterations are required, the augmentation acts as a global correction of the successive search directions (see figure (1)). It ensures that any solution of the form (40) still satisfies the reduced balance equations without local/global enhancement. Therefore, the complementary solution $\delta \underline{\mathbf{U}}_{\text{K}}^{(f)}$ can be interpreted as a local correction to the solution provided by the pre-computed reduced model. We will see in the examples that, as the conjugate gradient does not need to search for the part of the solution that is directly captured in the pre-computed reduced space, the convergence rate is very high.

The diagonal preconditioner $\widetilde{\underline{\mathbf{M}}}^{-1}$ classically accounts for local heterogeneities in the structure. A much more sophisticated preconditioner can of course be used (e.g.:incomplete Cholesky factorization). Yet, if the reduced basis is correctly pre-computed and/or updated, the computational effort required to obtain a good solution of the projected problem should be relatively low. Hence, the potentially expensive construction of an advanced preconditioner is not *a priori* justified. The example section will show that the number of conjugate gradient iterations required to compute a correction to the

reduced model is indeed kept very low, even while only using this crude preconditioner.

4 Application to the adaptive reduction of damage problems

4.1 Damageable lattice model

We focus on a very simple damageable lattice structure, made of bars under traction or compression (figure (3)). Elementary bar indexed b occupies domain Ω_b such that $\Omega = \bigcup_{b \in \llbracket 1, n_b \rrbracket} \Omega_b$, where n_b is the number of bars. The direction of bar number b is denoted by $\underline{n}_{kl} = P_k P_l / \|P_k P_l\|$ where P_k and P_l ($k < l$ and $(k, l) \in \llbracket 1, n_P \rrbracket^2$) are the two extremities of the bar as defined in figure (2). The displacement field in is supposed constant in a cross-section. Therefore, the displacement of any point $M \in \Omega_b$ can be expressed in the following form:

$$\underline{u}|_{M \in \Omega_b} = (\underline{u}(\tilde{x}) \cdot \underline{n}_{kl}) \underline{n}_{kl} = u(\tilde{x}) \underline{n}_{kl} \quad (45)$$

where each cross section is parametrised by a scalar $\tilde{x} \in [0, \|P_k P_l\|]$, defined by $\tilde{x} = \underline{P}_k M \cdot \underline{n}_{kl}$.

We use a constitutive law based on classical damage mechanics to describe the behaviour of the lattice structure. The lineic strain energy of bar b reads:

$$e_{d|\tilde{x}} = \frac{1}{2} E (1 - d) S_b \epsilon_{|\tilde{x}}^2 \quad (46)$$

where $\epsilon_{|\tilde{x}} = u_{,\tilde{x}}(\tilde{x})$ is the strain measured in the direction of the bar, $(S_b)_{b \in \llbracket 1, n_b \rrbracket}$, is the section of the bar, and d is a damage variable which ranges from 0 (undamaged material) to 1 (completely damaged material). The local state equations, derived from the strain energy read:

$$\begin{cases} N_{|\tilde{x}} = \frac{\partial e_d}{\partial \epsilon} = E(1 - d) S_b \epsilon_{|\tilde{x}} \\ Y_{|\tilde{x}} = -\frac{\partial e_d}{\partial d} = \frac{1}{2} E S_b \epsilon_{|\tilde{x}}^2 \end{cases} \quad (47)$$

where we have introduced the axial force $N = S_b (\underline{n}_{kl}^T \underline{\sigma} \cdot \underline{n}_{kl})$, and the thermodynamic force Y . A non-local evolution law is defined to link the damage variables to the thermodynamic forces:

$$d_{|\tilde{x}} = \max_{\tilde{x} \in [0, \|P_k P_l\|]} \left(\min \left(1, \max_{\tau \leq t} (\alpha (Y_{|\tilde{x}, \tau})^\beta) \right) \right) \quad (48)$$

We suppose that the lattice is not subjected to any volume force, that the material of each elementary bar is isotropic and homogeneous, and that the section of each bar is constant. As a consequence, the resulting displacement field evolves linearly along the direction of each bar.

In terms of finite element discretization, each bar will be considered as a linear element

$$\underline{u}(\tilde{x}) = \left(1 - \frac{\tilde{x}}{\|P_k P_l\|} \right) \underline{u}|_{P_l} + \frac{\tilde{x}}{\|P_k P_l\|} \underline{u}|_{P_k} \quad (49)$$

Because of the assumptions made previously on the loading, geometry and material properties, the finite element solution yields the “exact” solution to the damageable lattice problem.

4.2 Lattice structure and loading

We solve the problem defined in figures (3). The sections and Young’s modulus are unitary. The material parameters are $\beta = 0.5$ and $\alpha = \sqrt{2}$. Any bar parallel to axis (O, \underline{x}) , (O, \underline{y}) or (O, \underline{z}) has a unitary length.

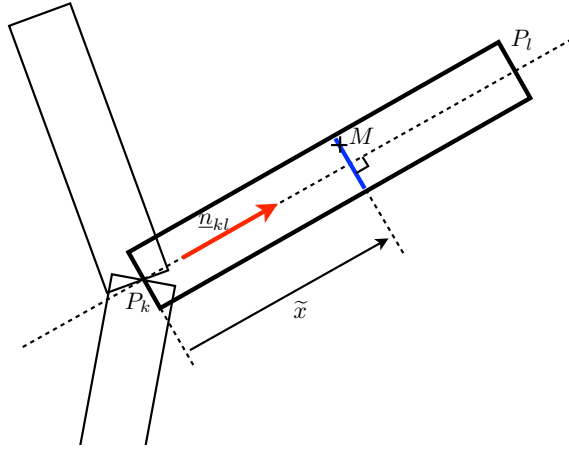


Figure 2: Definition of the lattice problem.

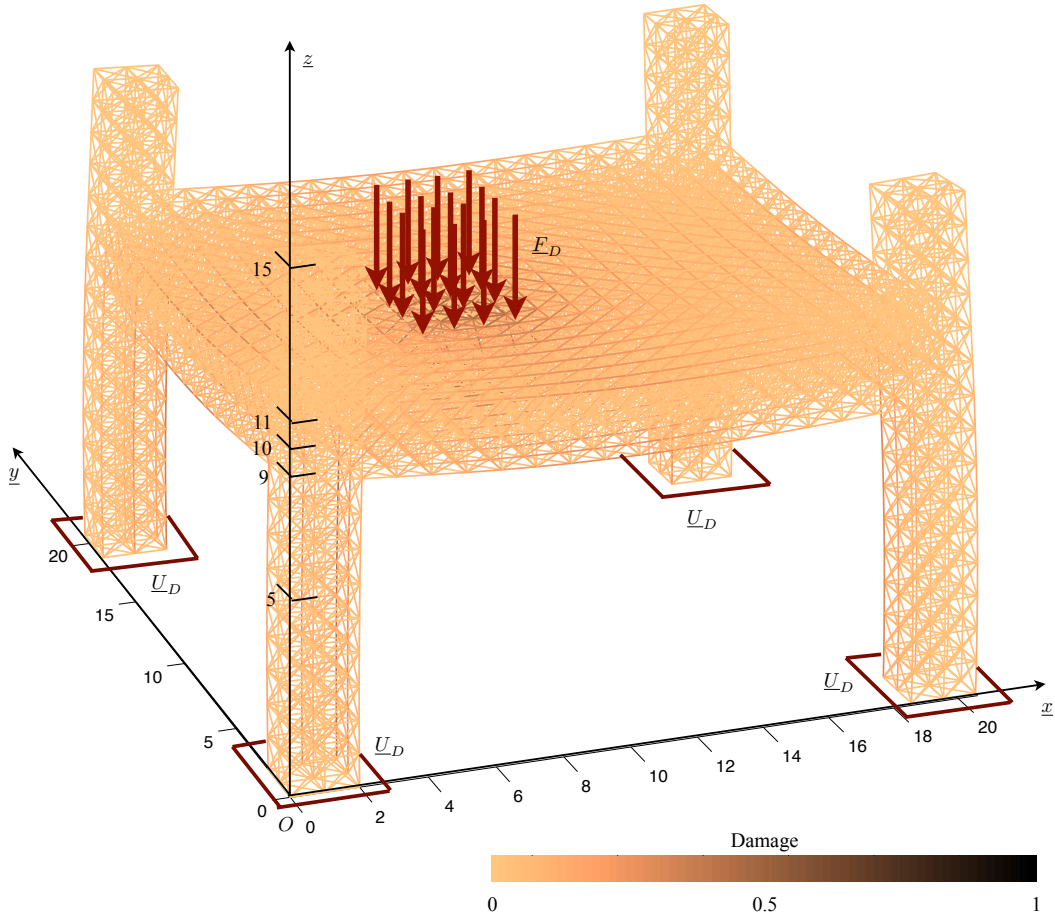


Figure 3: Reference solution obtained at the eighth load increment of the full-scale simulation (maximum value of prescribed forces before unstable damage propagation). Darker bars have a higher damage state.

The Dirichlet boundary conditions applied to the structure read, for any node (P_i) (with $i \in \llbracket 1, n_p \rrbracket$) of the lattice such that $(P_i) \in (O, \underline{x}, \underline{y})$:

$$\underline{u}_{|P_i} \cdot \underline{z} = 0 \quad (50)$$

As represented in figure (3), homogenous Neuman boundary conditions, in the $-\underline{z}$ direction, are applied to any point P_i that belongs to plane \mathcal{P} defined by:

$$\mathcal{P} = \{M = (x, y, z) \mid x \in [7, 9] \text{ and } y \in [8, 10] \text{ and } z = 11\} \quad (51)$$

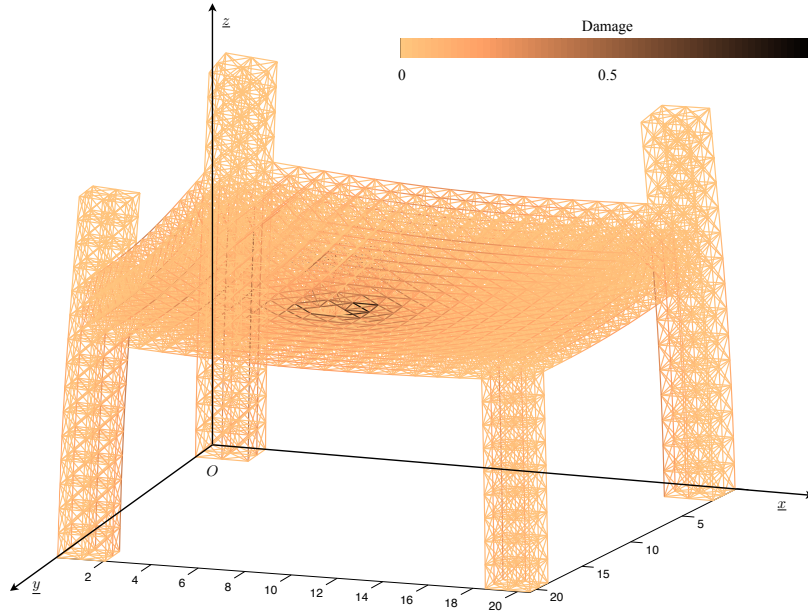


Figure 4: Final damage obtained at the end of the full-scale simulation, performed in 30 load increments. The bars that are completely damaged are not represented.

Increasing the value of the applied forces leads to the initiation and propagation of a damaged zone as shown in figure (4). This phenomenon is unstable from a structural point of view, and the damage state represented in this figure is past the limit point at which the force applied to the structure reaches its maximum value. An arc-length algorithm with local control is used to solve the damage propagation beyond this limit point. The basic idea of this algorithm is to look for the amplitude of the force under the constraint that the maximum increment in the damage variables over a time step is fixed (see [45, 2] or [18] for additional details). The reference solution is obtained by solving the full problem (6) in 30 of these damage increments of the Newton/arc-length algorithm. The final state is represented in figure (4).

Our purpose is to check the adequacy of the proposed local/global reduction technique when trying to obtain at cheap costs the load/deflection curve (pre and post failure phases) of the structure under the loading conditions described previously. The results provided in the following sections give the value of the norm of the external forces vector (also called loading factor) obtained by the arc-length method. The displacement used to plot the load/deflection curves is the average vertical displacement of the points at which forces are applied.

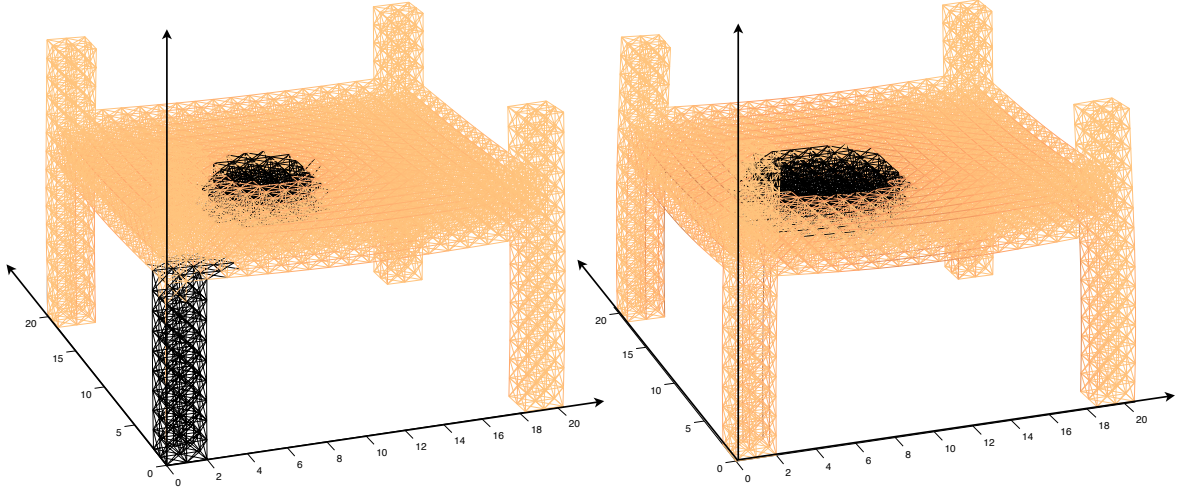


Figure 5: Splitting obtained at two different stages of the simulation performed by the local/global reduction method (left: load increment 4, right: load increment 30). Dark bars are connected to at least one fully resolved node. As time increases, the degrees of freedom for which no reduction is performed localises in the region of failure.

4.3 Particularisation of the local/global approach

4.3.1 Definition of the Local/Global splitting

We need at that point to particularize the general method developed in section 3. The splitting between reduced degrees of freedom $\widetilde{\Delta \mathbf{U}}^{(r)}$ and the complementary ones $\widetilde{\Delta \mathbf{U}}^{(f)}$ is made by considering that the part of the structure undergoing strong damage variations will not be correctly solved for when projected on a pre-computed global basis (see [38, 37] for similar assumptions in a domain decomposition-based local/global approaches). This statement is purely based on mechanical intuition and will be validated in the following examples.

The following procedure is adopted. The splitting will be fixed over a time step, to avoid the stagnation or divergence of the Newton process. At the end of a time step, the element undergoing the maximum damage increment is spotted, and a sphere of radius ρ_s , centred at the isobarycenter of the element is created. Every degree of freedom belonging to a node located inside this sphere is set as a fully resolved degree of freedom.

This procedure is repeated on the remaining elements (those which do not have all their degrees of freedom in the previous list).

The algorithm is stopped if one of the two following statements is satisfied.

- The maximum damage increment in the remaining elements is lower than k_{Dam} times the global maximum. In our examples, k_{Dam} is set to 0.5. This ensures that the fully resolved degrees of freedom are actually connected to elements undergoing a significant increase in their damage state compared to the remaining of the structure.
- The number of fully resolved degrees of freedom exceeds k_{LocGlo} times the total number of unknowns. k_{LocGlo} is set to 0.1 in our examples. This statement will be satisfied if the former one is not, which means that the structure undergoes a virtually homogeneous increase in its damage state. In that case, the local/global procedure might be inefficient.

To illustrate this procedure, two different reduced domains are represented in figure (5). The dark elements are connected to at least one fully resolved node. The first one is obtained at the end of the

third time step and is used to solve the fourth time step of the analysis, while the second one is used to solve the thirtieth and last time step. The first figure is obtained during the initiation phase. Notice that the maximum damage increments are found in the zone where the load is applied, and in the “pillar” that is closer to this zone, which undergoes a significant bending loading. In the later case, corresponding to unstable propagation phase, the damage increment localises in the zone where the load is applied.

This empirical choice for the local/global splitting is only justified by its efficiency. However we only considered one type of model for brittle fracture. We expect the optimal splitting to depend on the chosen model and on the stage of the structural failure. A more suitable criterion should be investigated in future work, based on algebra to circumvent this dependency. For instance, one could consider the decrease rate of the singular values of the correlation matrix restricted to an assumed smoothed domain and loosely maximise it, with a greedy-like algorithm, by successive small modifications of this domain.

4.3.2 Modification of the reduced basis

At the end of a time step, the reduced basis is enriched in order to take into account the information provided by the local iterative solver. The procedure used here is simply to add the successive solutions to the initial snapshot, and sort them by a singular value computation. This step can be optimized, following our previous work [18]. The issue here is that the reduced basis is only locally enhanced, which might results in ill-posed reduced problems if not correctly addressed. This question will not be further discussed in this paper, which focuses on the benefits provided by the local/global approach.

4.4 *A posteriori* reduced basis

4.4.1 Reference solution to construct the *a posteriori* reduced basis

The first test consists in using as a snapshot the consecutive reference solutions obtained by a Newton/arc-length solution algorithm performed on the initial system of equations, without reduction. In other terms, we try to reproduce the snapshot by reduced order modelling. In terms of realistic applications, this example is not of particular interest as one would need to know the solution in advance in order to obtain the following results. However, studying this academic case is interesting to understand the behaviour of the proposed strategy.

The snapshot is composed of $n_s = 30$ vectors. This information is compressed by using a singular value decomposition as detailed in section 2.2.2. We plot, in figure (6), the normalised error in the snapshot representation

$$\nu_{\text{SVD}} = \frac{\left(\sum_{j=1}^{n_s} \left\| \mathbf{s}^j - \sum_{i=1}^{n_c} (\mathbf{C}^{iT} \mathbf{s}^j) \mathbf{C}^i \right\|^2 \right)^{\frac{1}{2}}}{\left(\sum_{j=1}^{n_s} \|\mathbf{s}^j\|^2 \right)^{\frac{1}{2}}} = \frac{\left(\sum_{i=n_c+1}^{n_s} \lambda_i \right)^{\frac{1}{2}}}{\left(\sum_{i=1}^{n_s} \lambda_i \right)^{\frac{1}{2}}} \quad (52)$$

as a function of the order of truncation of the SVD, where the eigenvalues $(\lambda_i)_{i \in \llbracket 1, n_s \rrbracket}$ of the correlation matrix are now sorted in decreasing order. The decrease rate of this criterion is relatively low, which suggests that the information from the snapshot is poorly suited to compression. For reference, in [15], the author estimates that the snapshot space is correctly approximated by the space spanned by the first n_c singular vectors if $\nu_{\text{SVD}} \leq 10^{-8}$. In our case, this would lead to the selection of the 30 singular vectors to define a relevant reduced model. The singular vector basis is here truncated at order $n_c = 3$ in order to yield a significant (thus observable) error in the solution when using the basic POD method. This error can then be compared to the one obtained when using the proposed local/global scheme.

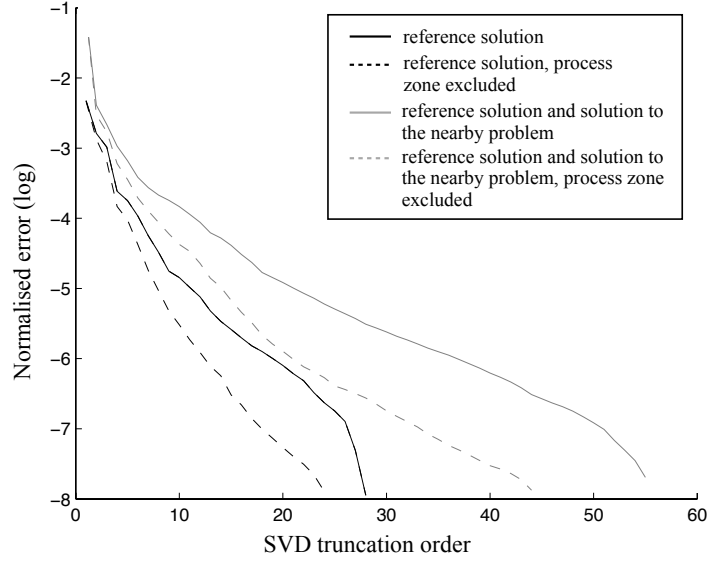


Figure 6: Measure of the error in the snapshot representation as a function of the error of truncation of the singular value decomposition. In solid lines dark line, the set of 30 reference solutions are defined as the snapshot. In grey lines, a second snapshot operator is defined by concatenating the 30 reference solutions and the 30 solutions of the nearby problem. Both analyses are repeated when excluding the process zone (defined *a priori* by the set of points $\{M = (x, y, z) \mid x \in [5, 13] \text{ and } y \in [6, 13]\}$) from the snapshot definition (i.e.: the snapshot now is the restriction of the solutions to the part of the domain which does not undergo significant damage). The resulting error curves are plotted in dashed lines.

In figure (7), the reference load/deflection curved is compared to the one obtained when using the basic POD on the first hand, and when using the local/global reduction technique on the other hand. The peak load obtained with the basic POD is slightly underestimated, by less than 2%, but the overall behaviour of the structure is correctly captured. The local/global algorithm permits to further increase the accuracy of the approximated solution. The error in the maximum load factor is now less than 0.5 %.

We now analyse the effect of preconditioning the local iterations of the conjugate gradient by the projection technique detailed in section 3.4.2. The convergence curves that are plotted in figure (8, left) are obtained at the third Newton iteration of the fiftieth time step of the analysis (i.e.: during the damage propagation phase). The norm of the condensed system (35) of equation, normalized by the norm of the condensed right-hand term is observed as a function of the number of iterations of the conjugate gradient solver. One can see that the number of iteration to convergence is significantly decreased when using the augmented Krylov solver.

An other interesting observation is that the convergence curve decreases in a monotonic manner, which means that the error in the global residual keeps decreasing as a function of the cost of the local conjugate gradient. This is not, of course, a proof, and this conclusion is probably problem dependent, but the same behaviour as been observed in all our numerical experiments. A probable reason for this is that the augmentation deflates the spectrum of the condensed operator, removing its highest eigenvalues and forcing the Krylov algorithm to its super-convergence phase (one can refer to [46] for a proper analysis of the effect of eigenvectors-based augmentation on the convergence of Krylov solvers).

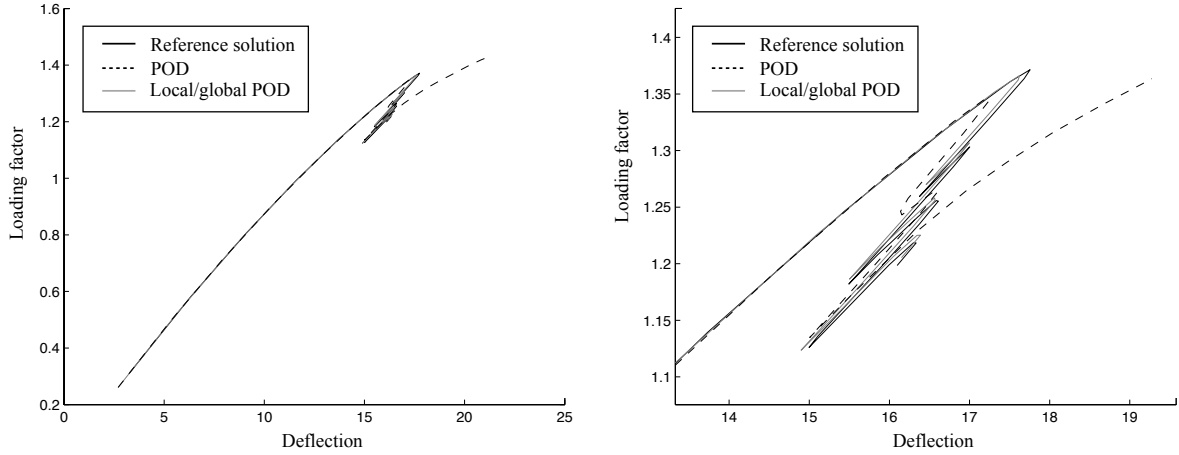


Figure 7: Load/deflection curves obtained one using successively direct numerical simulation, simple proper orthogonal decomposition and local/global reduction technique. The snapshot used here is the solution of the direct numerical simulation.

4.4.2 Nearby solution to construct the *a posteriori* reduced basis

In the second test case, we use as a snapshot the consecutive solutions obtained by solving a nearby problem, which differs from the reference one only by the position of the applied forces: homogenous Neumann boundary conditions are applied to any point P_i (with $i \in \llbracket 1, n_P \rrbracket$) such that $P_i \in \mathcal{P}'$ defined by:

$$\mathcal{P}' = \{M = (x, y, z) \mid x \in [9, 11] \text{ and } y \in [9, 11] \text{ and } z = 11\} \quad (53)$$

This snapshot is for instance a particular output of a parametric study meant to obtain the response of the system subjected to various load cases. One would expect to be able to reuse the information generated during this simulation to fasten the solution process of the next one. The difference between the snapshot solution and the reference solution that is looked for is illustrated in figure (9). It should be noted that far away from the part of the structure where Neumann boundary conditions are applied, the solutions are qualitatively similar. Yet, they are very different within this process zone, which qualitatively justifies the local/global splitting proposed in section (4.3.1). The same effect appears in figure (6, which shows that the combination of the information from the two load cases can be significantly compressed when removing the process zone from the SVD analysis.

Figure (10) compares the load/deflection curves obtained in the reference case, and when using successively the basic POD strategy and the local/global reduction technique. Obviously, the POD strategy loses its relevance very quickly, as the damage localises in the part of the structure where the damage would localise during the snapshot simulation. The resulting model is too stiff, and the maximum strength of the structure is overestimated. When using the local/global approach, with the parameters described previously for the fine analysis of the damaged areas, a much more accurate load/deflection curve is obtained. Yet the peak load is still overestimated by 5%. This remaining inaccuracy is due to the fact that the information from the snapshot is not correct in the reduced zone. Global corrections of the reduced basis are still necessary. This issue is discussed in the next section.

In figure (8, right) we analyse the effect of preconditioning the local iterations of the conjugate gradient by the projection technique. As in the previous test case, the convergence curves are observed at the third Newton iteration of the fiftieth time step of the analysis. Again, the number of iteration to convergence is significantly decreased when using the augmentation. We would expect this preconditioning technique to be less efficient in this particular case. Yet it is not the case, which is probably due to the fact that the global update described in section 4.3.2 is mechanically relevant.

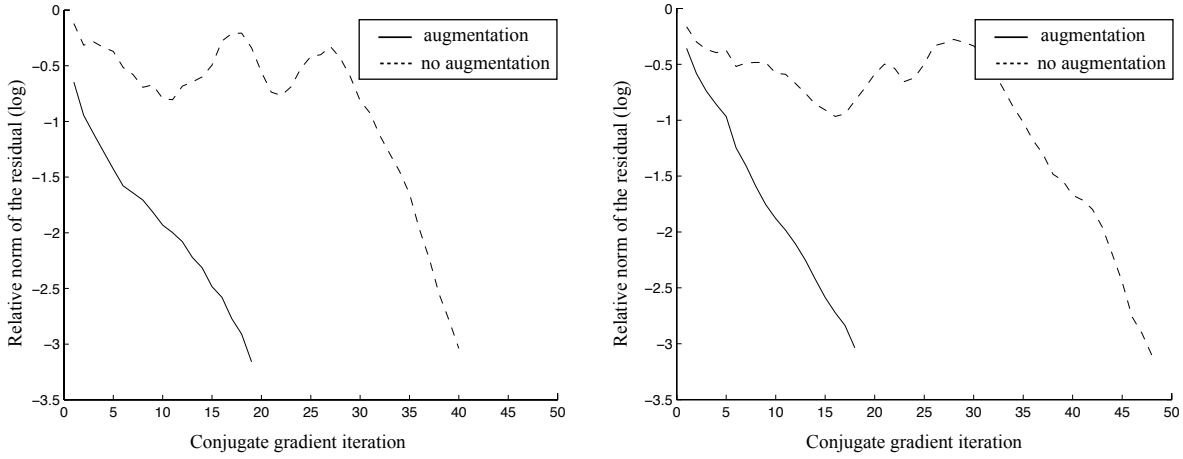


Figure 8: Convergence curves of the augmented conjugate gradient used to solve the local problem, at the third Newton iteration of load increment 15. On the left-hand side, the reference solution is used as a snapshot. On the right-hand side, the snapshot used to define the reduced model is the solution of the nearby problem.

5 Adaptive reduced basis

We now try to reduce the inaccuracy observed in figure (10) by using global corrections of the reduced model. The adaptation algorithm used here was introduced in [18]. We just recall the basics of this particular strategy, and explain how it can be used to enhance the proposed local/global scheme.

5.1 "On-the-fly" global corrections of the reduced basis

The principle of these global corrections, is to update the reduced basis used to define reduced problem (10). It is, for now, not linked to the global/local technique described previously in this paper. The updates are done "on-the-fly", at any iteration $i + 1$ of the Newton algorithm used to solve (10) at a given time step t_{n+1} of the analysis. If the reduced problem is sufficiently converged, and if the residual of the norm of the initial problem evaluated at iteration i is estimated too high, then the following linearized system is considered:

$$\underline{\underline{\mathbf{K}}}_T^i \delta \bar{\mathbf{U}} = -\underline{\mathbf{R}}^i \quad (54)$$

$\underline{\underline{\mathbf{K}}}_T^i$ is an approximation of the tangent $\underline{\mathbf{K}}_T^i$, while $\underline{\mathbf{R}}^i$ is the residual of the initial system of equations (6) evaluated in $\underline{\mathbf{U}}|_{t_n} + \underline{\underline{\mathbf{C}}}^i \underline{\alpha}^i$. Operator $\underline{\underline{\mathbf{C}}}^i$ designates the reduced basis considered at iteration i of the Newton process.

The solution to this problem is searched for in two supplementary spaces by a projected Krylov algorithm:

$$\begin{aligned} \delta \bar{\mathbf{U}} &= \delta \bar{\mathbf{U}}_C + \delta \bar{\mathbf{U}}_K \\ \text{where } \begin{cases} \delta \bar{\mathbf{U}}_C \in \text{Im}(\underline{\underline{\mathbf{C}}}^i) \\ \delta \bar{\mathbf{U}}_K \in \text{Im}(\underline{\underline{\mathbf{C}}}^i)^\perp \mathbb{K}_T^i = \text{Ker}(\underline{\underline{\mathbf{C}}}^i{}^T \underline{\underline{\mathbf{K}}}_T^i) \end{cases} \end{aligned} \quad (55)$$

This problem is solved coarsely by a projected Krylov solver (10^{-1} is the typical value of the stopping criterion), orthogonally to the current reduced basis. The resulting solution, which is $\underline{\underline{\mathbf{K}}}_T^i$ -orthogonal to the $\underline{\underline{\mathbf{C}}}^i$, is added to the reduced basis:

$$\underline{\underline{\mathbf{C}}}^{i+1} = \left(\underline{\underline{\mathbf{C}}}^i \quad \frac{\delta \bar{\mathbf{U}}_K}{\|\delta \bar{\mathbf{U}}_K\|} \right) \quad (56)$$

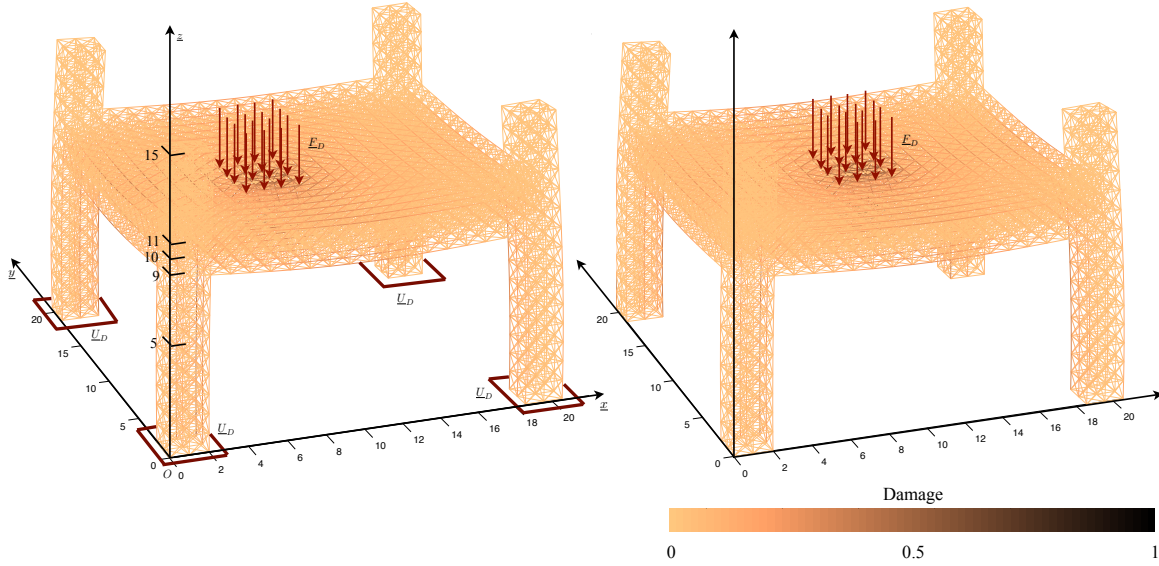


Figure 9: Two set of reference solutions obtained by direct numerical simulations and used as snapshot vectors in our examples. Only the solutions to the eighth load increment are represented in both cases. On the left-hand side, the snapshot is the reference solution which is looked for. On the right-hand side, the snapshot is the solution to a slightly different problem (“nearby problem”), where the position of the applied load has been modified.

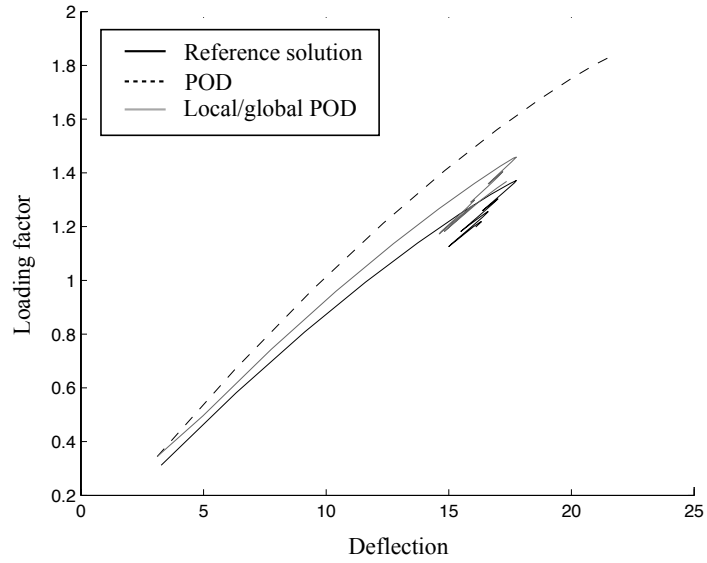


Figure 10: Load/deflection curves obtained one using successively direct numerical simulation, simple proper orthogonal decomposition and local/global reduction technique. The snapshot used here is the solution of the direct numerical simulation to the nearby problem.

Newton prediction $i + 1$ can now be solved after updating residual $\underline{\mathbf{R}}_{\mathbf{R}}^i$ and reduced stiffness $\underline{\mathbf{K}}_{\mathbf{T},\mathbf{R}}^i$ with respect to this new basis vector.

Various features are added to this algorithm to efficiently sort the successively added basis vectors, but the main idea described above is sufficient to understand the following results.

5.2 Coupling with the local/global approach

Once the global norm of the residual of locally reduced problem (27) is sufficiently converged, and if the global residual is judged too high, a correction vector to the global reduced basis is calculated, following the algorithm given in the previous subsection. The residual and tangent operator of problem (27) are updated and condensed tangent problem (35) is solved with the projected conjugate gradient algorithm. This algorithm converges in a very few iterations as the new basis vector, added to the augmentation space, is very close to the solution which is looked for. The slight mismatch is only due to:

- the approximation of the tangent used to provide the global correction vector.
- the coarse solution of (54).
- the approximation of the displacement in the reduced zone of the local/global approach.

5.3 Results

The test case described in section 4.4.2 is solved using successively the local/global technique on its own, the global correction algorithm on its own, and the local/global reduction technique coupled with global corrections. The target norm of the relative global residual norm is 10^{-1} . Corrections are allowed only if the relative norm of the residual of locally reduced problem (27) is below 10^{-3} . It is clear from the load-displacement curve in Figure 11 that the global corrections significantly improve the results of the local/global algorithm. (11). An illustration of the POD basis and its successive corrections is given in figure (13).

More importantly, the number of relatively expensive global corrections required to achieve the target in terms of global residual norm is very low when using the local/global approach, as opposed to the one observed when using only global corrections. This effect is shown in figure (12). One can also notice that using only the global correction methods lead to an increase in the computational costs in the damage propagation phase of the analysis, which has been reported in [18]. With the local/global reduction approach, this effect disappears. The same trends are observed in [37], where a local/global algorithm was developed in a domain decomposition framework to avoid unnecessary computations far away from the process zones. This suggests that the result obtained here may be extended, and that the number of global corrections required to achieve a given level of accuracy does not depend on the local nonlinearity (i.e: independent on the damage state).

6 Conclusions and discussion

The work described in this paper focuses on decreasing the computational effort required in solving large scale damage and fracture problems, where small scale phenomena must be taken into account to accurately represent the global structural behaviour. Although most investigations on this issue rely on the extension of homogenization frameworks, we followed an alternative route, choosing projection-based model order reduction as a starting point. The proposed local/global reduction technique was demonstrated on a damaging 3D frame structure, but we expect that the algorithm will carry forward to the more general case of the initiation and propagation of cracks in structures.

The main conclusions of the paper are:

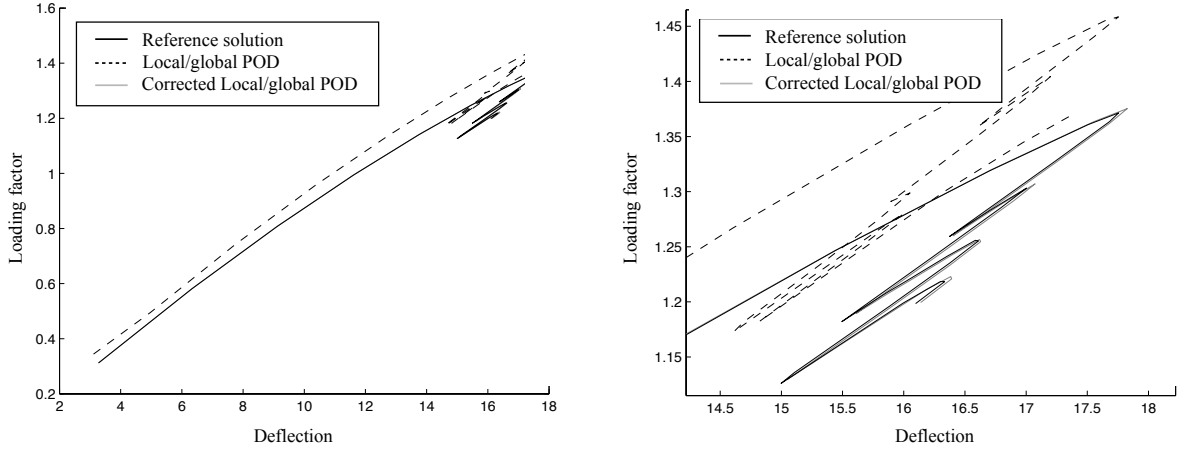


Figure 11: Load/deflection curves obtained one using successively direct numerical simulation, local/global reduction technique and local/global reduction technique with "on-the-fly" global corrections.

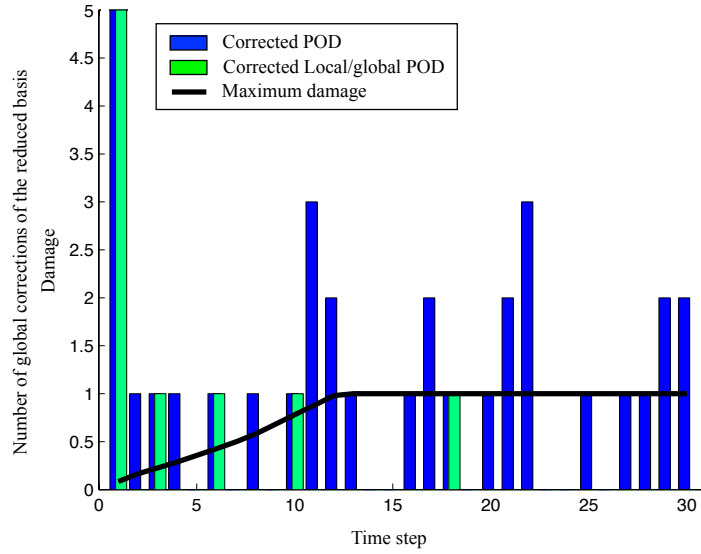


Figure 12: Number of global corrections performed when using the adaptive model order reduction on the one hand, and when using adaptive local/global model order reduction on the other hand.

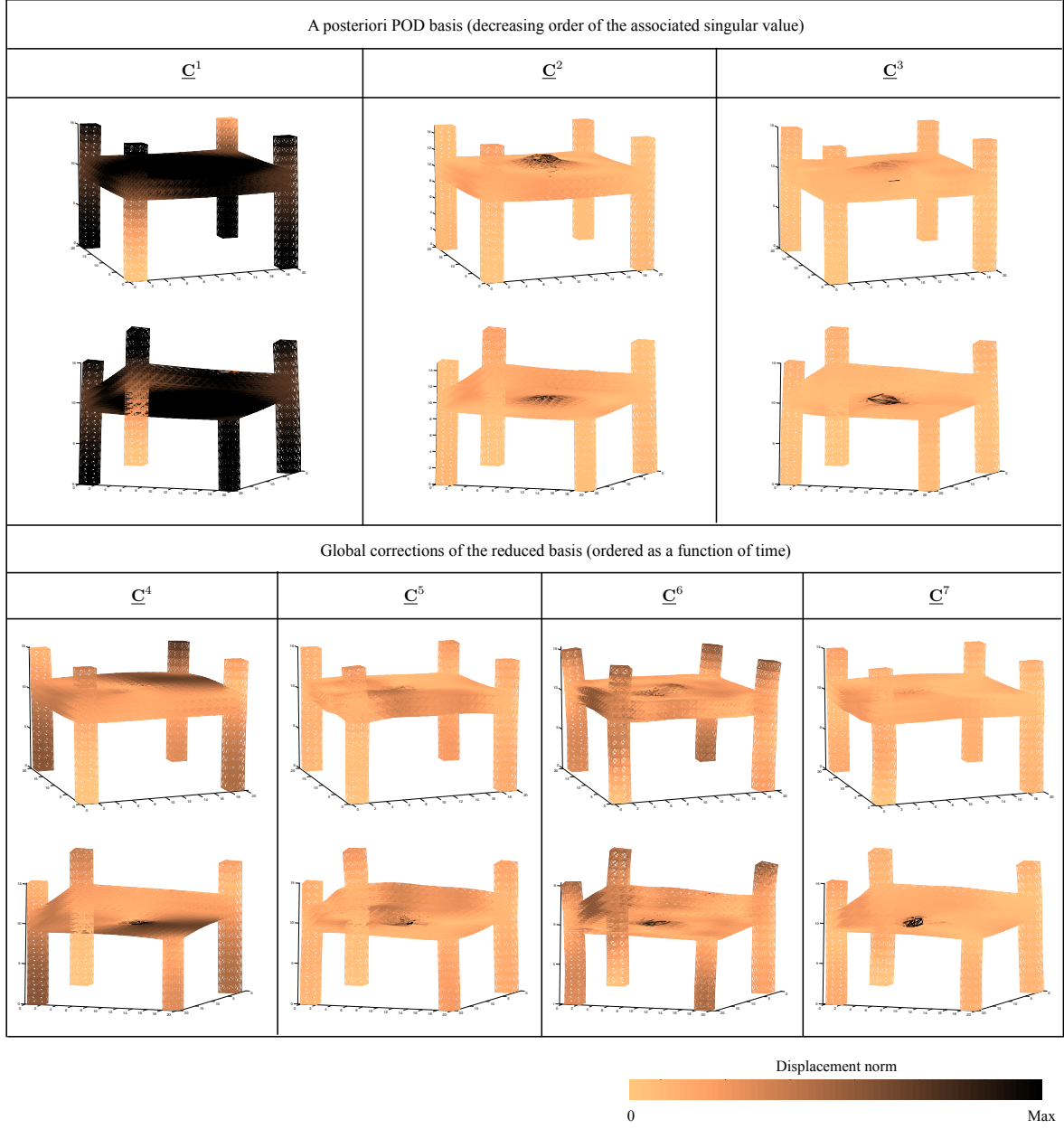


Figure 13: A posteriori POD basis (3 vectors) and successive global corrections (4 vectors computed respectively at time steps 1, 3, 10 and 18). Two different views of each basis vector are represented for better qualitative understanding.

- The proposed local/global model order reduction approach can significantly improve the relevancy of using a global reduced basis to approximate the displacement field in the case of localised failure;
- The number of global corrections of the reduced model required to obtain a given level of accuracy is drastically reduced when the balance equations which exhibit the highest level of nonlinearity are excluded from the reduction.

In most of the applications for which this methodology is intended, the additional computational costs are kept minimal for two reasons:

- The damaged zone where no reduction is performed is small compared to the size of the structure in realistic engineering problems;
- The local fine scale solution is efficiently obtained by Krylov corrections of a good initialisation provided by the restriction of the reduced model to the damaged zone.

Yet, some issues need to be addressed before global/local reduction algorithms can be efficiently and safely used to tackle engineering problem:

- Cost of the construction of the reduced systems. A system reduction is needed to ensure the computational efficiency of the method. It is now well known that projection-based model order reduction only reduces the size of the algebraic systems to solve, and that the construction of these systems is still expensive. Indeed, it involves the evaluation of the internal forces to construct the tangent stiffness and the residual of the reduced problem, which requires global updates of the constitutive behaviour, and integrations over the whole domain. System approximations permit to solve this issue. They basically consist in obtaining a cheap, but reliable approximation of the internal forces. In [18], we used the so-called hyperreduction [15] (also introduced as the missing-point approximation by [47] in a different context). We found that adaptive reduced basis methods need to be carefully tailored when used within this framework. Indeed, one cannot rely on the fact that the residual of the full-scale problem is known at any iteration of the nonlinear solver anymore: the number of global evaluation of the internal forces must be kept minimal. In addition, we observed that the stability of the particular system approximation that we used is highly dependent on the choice of its parameters (reduced integration domain), especially when the damage mechanism localises. We foresee that excluding the most damaged zones from the reduced domain can significantly improve the stability of this scheme. Some numerical investigations are required to confirm this intuitive speculation.
- Quality control. Each step in the algorithm is a source of error: (i) choice of the set of degrees of freedom for which the solution is searched in a reduced space (ii) choice of the reduced basis (iii) choice of the accuracy for the local solution (iv) choice of the system reduction (v) choice of the time stepping parameters. These choices should be made within an encompassing framework to ensure that the error associated with each of them can be controlled in the same manner. Therefore, one needs to find a relevant error criterion and build a control process, capable of ensuring that the above choices introduce the same order of error in the solution so that the global error be not dominated by any of them in particular.

In terms of further applications, this method should be validated using more advanced models for brittle fracture, in particular nonlocal continuum damage or softening plasticity models, for which the distinction between local and global effects is not trivial. The extension to ductile failure and/or geometric nonlinearities should also be possible, the global corrections being expected to efficiently handle the “smooth” part of the nonlinearities, while the local/global scheme should treat separately the localisation of defects.

We believe that this local/global reduction framework is an important milestone to obtain an optimal reduction method applicable to the parallel simulation of failure in solids. Indeed, as alluded

to in the introduction of this paper, several authors have proposed to perform a systematic model order reduction of the expensive local problems (so-called “problems by substructure”) classically involved in domain decomposition methods [32, 33, 29]. The application of such ideas to failure assessment has never been tackled so far, as it first requires an efficient reduction scheme for the local problems, such as the one proposed in this paper, in order to guarantee reasonable scalabilities and load balancing.

7 Acknowledgements

Pierre Kerfriden would like to thank Pierre Gosselet (LMT/ENS Cachan, France) for his precious advices. Stéphane Bordas and Pierre Kerfriden would like to acknowledge the financial support of the Royal Academy of Engineering and of the Leverhulme Trust for Bordas’ Senior Research Fellowship “Towards the next generation surgical simulators” (2009-2010) as well as the support of EPSRC under grant EP/G042705/1 Increased Reliability for Industrially Relevant Automatic Crack Growth Simulation with the eXtended Finite Element Method. The support of the British Council Exchange Programme is also gratefully acknowledged by the authors.

References

- [1] N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal of Engineering Science*, 46:131–150, 1999.
- [2] O. Allix and A. Corigliano. Modeling and simulation of crack propagation in mixed-modes inter-laminar fracture specimens. *International Journal of Fracture*, 77:11–140, 1996.
- [3] A.D. Jefferson and T. Bennett. A model for cementitious composite materials based on micro-mechanical solutions and damage-contact theory. *Computers & Structures*, doi:10.1016/j.compstruc.2008.09.006, 2009.
- [4] P. Ladevèze and G. Lubineau. An enhanced mesomodel for laminates based on micromechanics. *Composites Science and Technology*, 62(4):533–541, 2002.
- [5] P. Grassl and M. Jirásek. Damage-plastic model for concrete failure. *International journal of solids and structures*, 43(22-23):7166–7196, 2006.
- [6] T. Rabczuk, G. Zi, S. Bordas, and H. Nguyen-Xuan. A geometrically non-linear three-dimensional cohesive crack method for reinforced concrete structures. *Engineering Fracture Mechanics*, 75(16):4740–4758, 2008.
- [7] J. Fish, K. Shek, M. Pandheeradi, and M. S. Shephard. Computational plasticity for composite structures based on mathematical homogenization: Theory and practice. *Computer Methods in Applied Mechanics and Engineering*, 148:53–73, 1997.
- [8] F. Feyel and J.-L. Chaboche. FE² multiscale approach for modelling the elastoviscoplastic behaviour of long fibre SiC/Ti composite materials. *Computer Methods in Applied Mechanics and Engineering*, 183(3-4):309–330, 2000.
- [9] Q.Z. Xiao and B.L. Karihaloo. Two-scale asymptotic homogenisation-based finite element analysis of composite materials. *Multiscale Modeling in Solid Mechanics: Computational Approaches*, page 43, 2009.
- [10] T. Belytschko, S. Loehnert, and J.-H. Song. Multiscale aggregating discontinuities: A method for circumventing loss of material stability. *International Journal for numerical Methods in Engineering*, 73:869–894, 2007.

- [11] V. P. Nguyen, O. Lloberas-Valls, Martijn Stroeve, and L. J. Sluys. Homogenization-based multiscale crack modelling: From micro-diffusive damage to macro-cracks. *Computer Methods in Applied Mechanics and Engineering*, doi:10.1016/j.cma.2010.10.013, 2010.
- [12] F. Cazès, M. Coret, A. Combescure, and A. Gravouil. A thermodynamic method for the construction of a cohesive law from a nonlocal damage model. *International Journal of Solids and Structures*, 46(6):1476–1490, 2009.
- [13] C. V. Verhoosel, J. C. Remmers, M. A. Gutiérrez, and R. De Borst. Computational homogenization for adhesive and cohesive failure in quasi-brittle solids. *International Journal for numerical Methods in Engineering*, doi:10.1002/nme.2854, 2010.
- [14] M. Barrault, Y. Maday, N.C. Nguyen, and A.T. Patera. An ‘empirical interpolation’ method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus de Mathématiques*, 339(9):667–672, 2004.
- [15] D. Ryckelynck. A priori hyperreduction method: an adaptive approach. *Journal of Computational Physics*, 202(1):346–366, 2005.
- [16] S. Niroomandi, I. Alfaro, E. Cueto, and F. Chinesta. Model order reduction for hyperelastic materials. *International Journal for numerical Methods in Engineering*, doi:10.1002/nme.2733, 2009.
- [17] N. Verdon, C. Allery and. C. Béghein, A. Hamdouni, and D. Ryckelynck. Reduced-order modelling for solving linear and non-linear equations. *International Journal for Numerical Methods in Biomedical Engineering*, doi:10.1002/cnm.1286, 2009.
- [18] P. Kerfriden, P. Gosselet, S. Adhikari, and S. Bordas. Bridging proper orthogonal decomposition methods and augmented newton-krylov algorithms: an adaptive model order reduction for highly nonlinear mechanical problems. *Computer Methods in Applied Mechanics and Engineering*, 200(5-8):850–866, 2010.
- [19] C. Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: an overview. *International Journal of Applied Mathematics and Computer Science*, 11(5):1093–1121, 2001.
- [20] J.W.S. Rayleigh and R.B. Lindsay. *The Theory of Sound*. Dover Publications, 1877.
- [21] J.M. Dickens, J.M. Nakagawa, and M.J. Wittbrodt. A critique of mode acceleration and modal truncation augmentation methods for modal response analysis. *Computers and Structures*, 62(6):985–998, 1997.
- [22] R. Craig and M. Bampton. Coupling of substructures for dynamic analysis. *American Institute of Aeronautics and Astronautics*, 6(7):1313–1319, 1968.
- [23] G. Kerschen, M. Peeters, J. C. Golinval, and A.F. Vakakis. Nonlinear normal modes, part i: A useful framework for the structural dynamicist. *Mechanical Systems and Signal Processing*, 23(1):170–194, 2009.
- [24] L. Sirovich. Turbulence and the dynamics of coherent structures. part I: coherent structures. *Quarterly of Applied Mathematics*, 45:561–571, 1987.
- [25] K. Karhunen. Über lineare methoden in der wahrscheinlichkeitsrechnung. *Annales Academiæ Scientiarum Fennicæ Series A1, Mathematical Physics* 37, 37:1–79, 1947.
- [26] M. Loeve. *Probability Theory*. 3rd edition Van Nostrand, Princeton, 1963.

- [27] A. Nouy. A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 196(45-48):4521–4537, 2007.
- [28] F. Chinesta, A. Ammar, and E. Cueto. Recent advances and new challenges in the use of the proper generalized decomposition for solving multidimensional models. *Archives of Computational Methods in Engineering - State of the Art Reviews*, 17(4):327–350, 2010.
- [29] P. Ladevèze, J.C. Passieux, and D. Néron. The latin multiscale computational method and the proper generalized decomposition. *Computer Methods in Applied Mechanics and Engineering*, 199(21):1287–1296, 2009.
- [30] A. Nouy and O. Le Maître. Generalized spectral decomposition method for stochastic non linear problems. *Journal of Computational Physics*, doi:10.1016/j.jcp.2008.09.010, 2008.
- [31] F. Galland, A. Gravouil, E. Malvesin, and M. Rochette. A global model reduction approach for 3d fatigue crack growth with confined plasticity. *Computer Methods in Applied Mechanics and Engineering*, doi:10.1016/j.cma.2010.08.018, 2010.
- [32] D. Rixen. A dual craig-bampton method for dynamic substructuring. *Journal of Computational and Applied Mathematics*, 168:383–391, 2004.
- [33] D. Markovic, A. Ibrahimbegovic, and K. C. Park. Partitioning based reduced order modelling approach for transient analyses of large structures. *Engineering Computations*, 26:46–68, 2009.
- [34] A. Ammar, E. Prulière, J. Ferec, F. Chinesta, and E. Cueto. Coupling finite elements and reduced approximation bases. *European Journal of Computational Mechanics*, 18(5-6):445–463, 2009.
- [35] D. Ryckelynck, F. Chinesta, E. Cueto, and A. Ammar. On the a priori model reduction: overview and recent developments. *Archives of Computational Methods in Engineering, State of the Art Reviews*, 13(1):91–128, 2006.
- [36] C. Farhat and F.-X. Roux. Implicit parallel processing in structural mechanics. *Computational Mechanics Advances*, 2:1–24, 1994.
- [37] P. Kerfriden, O. Allix, and P. Gosselet. A three-scale domain decomposition method for the 3D analysis of debonding in laminates. *Computational Mechanics*, 44(3):343–362, 2009.
- [38] L. Gendre, O. Allix, P. Gosselet, and F. Comte. Non-intrusive and exact global/local techniques for structural problems with local plasticity. *Computational Mechanics*, 44(2):233–245, 2009.
- [39] O. Allix, P. Kerfriden, and P. Gosselet. A relocation technique for the multiscale computation of delamination in composite structures. *Computer Modeling in Engineering and Sciences*, 55(3):271–291, 2010.
- [40] Y. Saad. *Iterative Methods for Sparse Linear Systems Iterative Methods for Sparse Linear Systems, Second Edition*. Society for Industrial and Applied Mathematics, 2003.
- [41] D.A. knoll and D.E. Keyes. Jacobian-free newton–krylov methods: a survey of approaches and applications. *Journal of Computational Physics*, 193:357–397, 2004.
- [42] R.S. Dembo, S.C. Eisenstat, and T. Steihaug. Inexact newton methods. *SIAM Journal on Numerical Analyses*, 19(2):400–408, 1982.
- [43] F. Risler and C. Rey. Iterative accelerating algorithms with krylov subspaces for the solution to large-scale nonlinear problems. *Numerical algorithms*, 23:1–30, 2000.

- [44] Z. Dostál. Conjugate gradient method with preconditioning by projector. *International Journal of Computer Mathematics*, 23(3):315–323, 1988.
- [45] E. Riks. The application of newton’s methods to the problem of the application of newton’s methods to the problem of elastic stability. *Journal of Applied Mechanics*, 39:1060–1065, 1972.
- [46] P. Gosselet. *Méthodes de décomposition de domaine et méthodes d’accélération pour les problèmes multichamps en mécanique non-linéaire*. PhD thesis, Université Paris 6, 2003.
- [47] P. Astrid, S. Weiland, K. Willcox, and A.C.P.M. Backx. Missing point estimation in models described by proper orthogonal decomposition. *IEEE Transactions on Automatic Control*, 53(10):2237–2251, 2008.